LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

numero 165
maggio 1982
anno XV
volume XXVIII

Alla ricerca dei bosoni vettori intermedi

È probabile che presto queste particelle pesanti, previste dalla teoria come portatrici della forza nucleare debole, possano venir rivelate fra i frammenti di collisioni tra protoni e antiprotoni

di David B. Cline, Carlo Rubbia e Simon van der Meer

Uno dei maggiori risultati conseguiti dalla fisica moderna negli ultimi 15 anni circa è stato lo sviluppo di una nuova classe di teorie unificate per la descrizione delle forze agenti tra le particelle elementari. Prima dell'introduzione di queste teorie le quattro forze osservabili della natura sembravano del tutto indipendenti una dall'altra. La forza elettromagnetica regola le interazioni di particelle elettricamente cariche; la forza nucleare debole è responsabile di processi quali il decadimento beta di un nucleo atomico radioattivo; la forza nucleare forte tiene assieme il nucleo, mentre la gravità tiene assieme l'universo. Quella di maggior successo tra le nuove teorie sancisce un legame tra la forza elettromagnetica e quella debole, suggerendo che esse siano semplicemente manifestazioni differenti di una sola forza fondamentale.

La teoria unificata elettrodebole sta per essere sottoposta a una verifica sperimentale decisiva. Una sua previsione fondamentale, infatti, è l'esistenza di tre particelle pesanti chiamate bosoni vettori intermedi (in inglese chiamati anche *weakon*, da *weak*, debole). Il primo acceleratore al mondo con energia sufficiente a creare tali particelle è stato ultimato da poco presso l'Organizzazione Europea per le Ricerche Nucleari (CERN) di Ginevra. Questo acceleratore, che era stato costruito originariamente per lanciare protoni di alta energia contro un bersaglio fisso, ora è stato adattato a un nuovo modo di funzionamento, nel quale protoni e antiprotoni si scontrano frontalmente.

Se i bosoni vettori intermedi esistono e se hanno le proprietà loro attribuite dalla teoria elettrodebole, dovremmo poter essere presto in grado di rivelarli. In questo momento costituiscono il trofeo più ambito di tutta la fisica e la loro scoperta porrebbe termine a una ricerca iniziata più di 40 anni fa.

Secondo l'interpretazione più accreditata delle interazioni tra particelle elementari, la trasmissione di una forza tra due particelle avviene con lo scambio di una terza particella intermedia. Una tale descrizione è l'essenza di una teoria quantistica dei campi. Per spiegare come le particelle possano interagire a distanza è necessario il concetto di un campo che si estende nello spazio; e si tratta di un campo quantistico perché è descritto da unità discrete, cioè le particelle intermedie. Nelle interazioni elettromagnetica e debole la particella scambiata è un membro della famiglia detta dei bosoni vettori. Questo termine si riferisce a una classificazione delle particelle secondo una delle loro più fondamentali proprietà: il momento angolare di spin. Un bosone (chiamato così dal nome del fisico indiano S. N. Bose) è una particella il cui spin, se misurato in unità fondamentali, è un intero, come 0, 1 o 2. Il termine «vettore» designa un bosone il cui valore di spin è uguale a 1.

Nel caso dell'elettromagnetismo il bosone vettore scambiato è il fotone, il «pacchetto d'onda» di energia elettromagnetica privo di massa e di carica che funge da quanto del campo elettromagnetico. I fotoni si osservano facilmente sperimentalmente (per esempio, sotto forma di luce) e studiandone le proprietà i fisici hanno costruito quella teoria estremamente precisa e comprensiva chiamata elettrodinamica quantistica, o QED, che è la teoria quantistica dei campi della forza elettromagnetica.

Il corrispondente portatore della forza nelle interazioni deboli è il bosone vettore intermedio (intermedio semplicemente per il suo ruolo di mediatore tra le particelle). L'esistenza di tale particella fu suggerita per la prima volta nel 1935 dal fisico giapponese Hideki Yukawa, che in quel tempo stava cercando una spiegazione unificata delle due forze nucleari da poco scoperte: quella forte e quella debole. Yukawa notò che il raggio d'azione di una forza doveva essere inversamente proporzionale alla massa della particella che la trasmette. Per esempio, il raggio d'azione della forza elettromagnetica è infinito, in accordo con il fatto che il fotone è privo di massa. D'altra parte le due forze nucleari hanno soltanto un raggio d'azione limitato; questo è il motivo per cui Yukawa pensò che

Fasci di protoni e di antiprotoni rotanti in senso opposto vengono generati al CERN come prodotti finali di una sequenza di eventi in più anelli acceleratori tra loro collegati. Dapprima un fascio di protoni *(in grigio)* viene accelerato fino a un'energia di 26 GeV (miliardi di elettronvolt) nel Protosincrotrone (PS). I protoni vengono poi indirizzati su un bersaglio metallico, producendo (tra l'altro) alcuni antiprotoni con un'energia di 3,5 GeV. Gli antiprotoni *(in colore)* vengono raccolti e trasferiti all'accumulatore di antiprotoni (AA) dove vengono combinati con antiprotoni iniettati precedentemente e concentrati in densi «fiotti» formati da centinaia di miliardi di particelle. I fiotti di antiprotoni vengono poi rinviati all'anello PS, dove vengono accelerati a 26 GeV. Gli antiprotoni da 26 GeV vengono iniettati nell'anello SPS, nel quale stanno già ruotando in senso opposto protoni con la stessa energia. Infine i due fasci vengono entrambi accelerati a 270 GeV nell'anello più grande. La pianta complessiva dell'installazione *(in basso)* mostra la posizione dei nuovi rivelatori di particelle, disposti in due lunghe sezioni rettilinee dell'anello SPS nelle quali i fasci circolanti in senso opposto vengono fatti collidere. Gli altri anelli mostrati vengono usati per vari altri esperimenti.

dovessero essere trasportate da particelle dotate di massa.

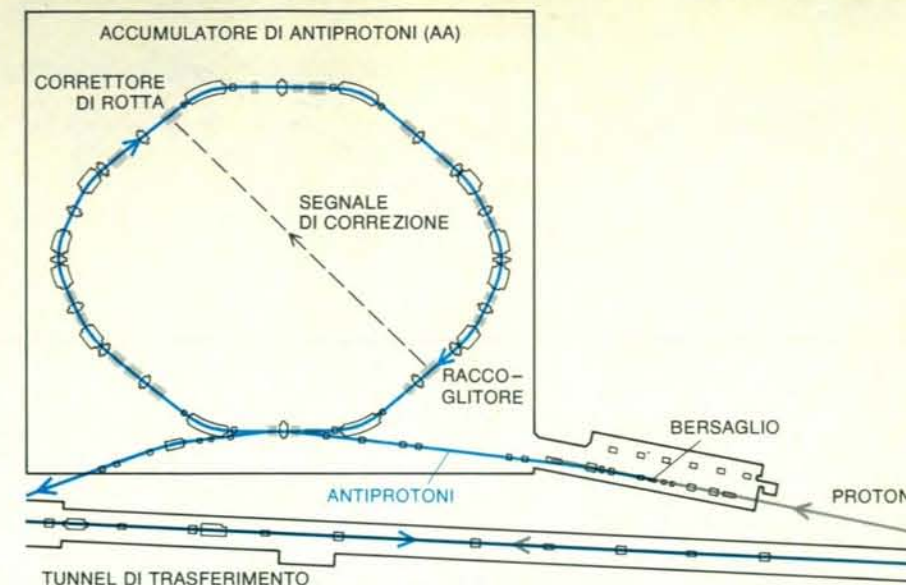
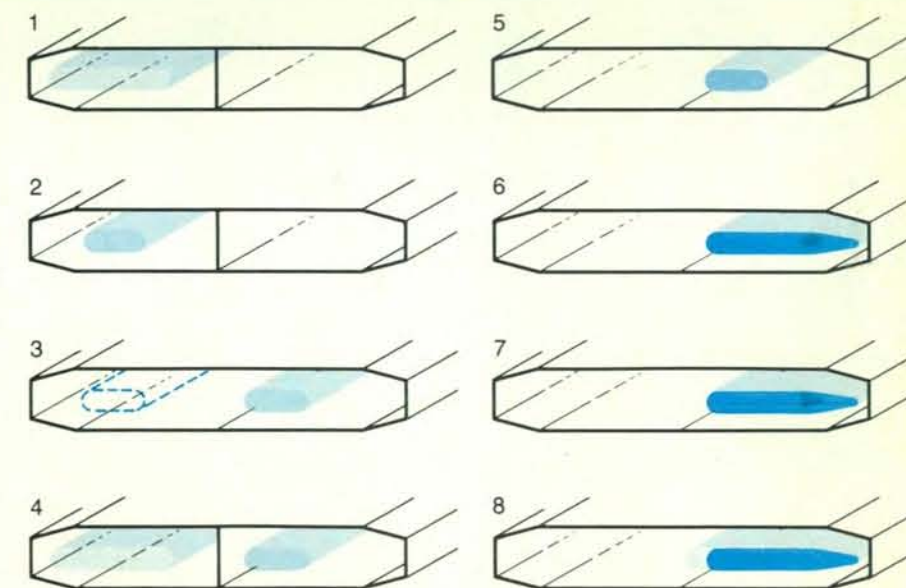
In particolare, Yukawa ipotizzò l'esistenza di una particella moderatamente pesante, chiamata in seguito mesone, il cui scambio dà origine alla forza attrattiva forte tra il protone e il neutrone. La prima particella di questo tipo che sia stata correttamente identificata, il mesone pi (o pione), fu scoperta nel 1947 nello sciame di particelle secondarie prodotte dalla collisione di una particella dei raggi cosmici con un atomo dell'atmosfera; oggi si possono produrre grandi quantità di mesoni a volontà con gli acceleratori di particelle.

Oggi si pensa che le particelle del nucleo, e tra esse i mesoni, siano formate da costituenti più fondamentali chiamati quark. I quark sono tenuti assieme dalla forza forte, ma in questo contesto la forza ha una forma del tutto differente da quella osservata tra protoni e neutroni. Si ritiene che la forza tra i quark sia trasmessa dalla famiglia di otto bosoni vettori privi di massa chiamati gluoni. Ai quark e ai gluoni viene assegnata una proprietà con il nome arbitrario «colore» che assume nelle interazioni forti lo stesso ruolo che, nelle interazioni elettromagnetiche viene assunto dalla carica elettrica. Per ricordare tale analogia la teoria quantistica dei campi della forza nucleare forte è chiamata cromodinamica quantistica, o QCD.

La forza nucleare debole ha un raggio d'azione ancor più breve di quella forte che agisce tra protoni e neutroni. Si può perciò prevedere che i bosoni vettori intermedi della forza debole abbiano una massa maggiore di quella del mesone pi. I primi tentativi di rivelare le particelle intermedie associate alla forza debole furono infruttuosi, forse perché la maggiore massa dei bosoni li poneva al di fuori della portata degli acceleratori di particelle esistenti. Tuttavia, fino alla formulazione della teoria unificata elettrodebole fra la fine degli anni sessanta e l'inizio degli anni settanta, non disponevamo di alcuna buona stima della massa delle particelle della forza debole.

La teoria elettrodebole fu sviluppata indipendentemente da Steven Weinberg della Harvard University e da Abdus Salam del Centro internazionale di fisica teorica di Trieste, con ulteriori contributi di Sheldon Lee Glashow di Harvard e di altri. La teoria, che si può oggi considerare come la spiegazione «standard» delle interazioni elettromagnetica e debole, fu la prima a fare previsioni specifiche e verificabili sulle proprietà dei bosoni vettori intermedi, tra le quali la massa. Inoltre, la teoria richiedeva l'esistenza di tre di tali particelle, con cariche elettriche +1 (la W^+), −1 (la W^-) e zero (la Z^0). La migliore stima attuale della massa dei bosoni vettori intermedi, espressa in funzione della loro energia equivalente, è di 79,5 GeV per la W^+ e la W^- e di 90 GeV per la Z^0. (L'abbreviazione GeV sta per gigaelettronvolt, o miliardi di elettronvolt; per avere un termine di paragone, la massa del protone equivale a poco meno di un GeV.)

L'idea centrale della teoria standard è che sia l'elettromagnetismo sia la forza debole derivino da una sola e più fondamentale proprietà della natura. A energie eccezionalmente elevate (abbastanza elevate perché particelle quali le W e la Z vengano prodotte tanto facilmente quanto i fotoni) gli eventi mediati dalle due forze dovrebbero essere indistinguibili. Questa unificazione teorica si realizza assegnando il fotone e i bosoni vettori intermedi alla stessa famiglia di quattro particelle. Alle energie oggi raggiungibili non vi sono dubbi che gli eventi elettromagnetici siano del tutto diversi da quelli deboli; inoltre, sembrerebbe improbabile una parentela fra il fotone e le particelle W e Z, dato che il fotone è privo di massa e le altre tre particelle sono tra le più pesanti di cui si conosca l'esistenza. La discrepanza viene spiegata nella teoria standard con la nozione di simmetria spezzata, che distingue le forze al diminuire dell'energia: qualcosa di vagamente simile, per aiutare l'intuizione, si può trovare nel passaggio di una sostanza attraverso varie fasi, al diminuire della temperatura.

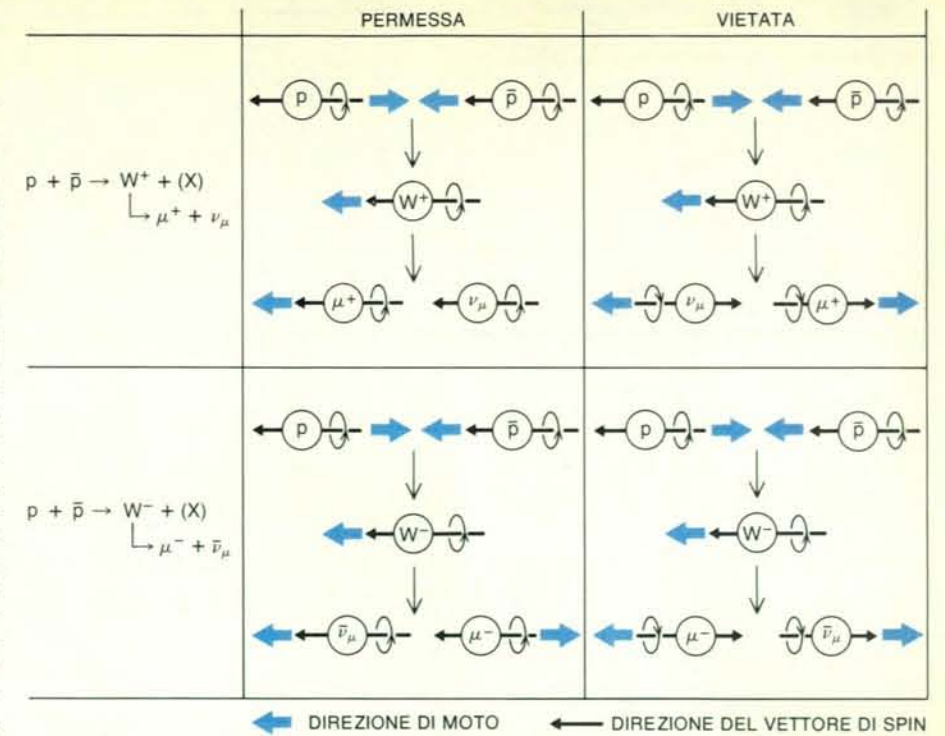
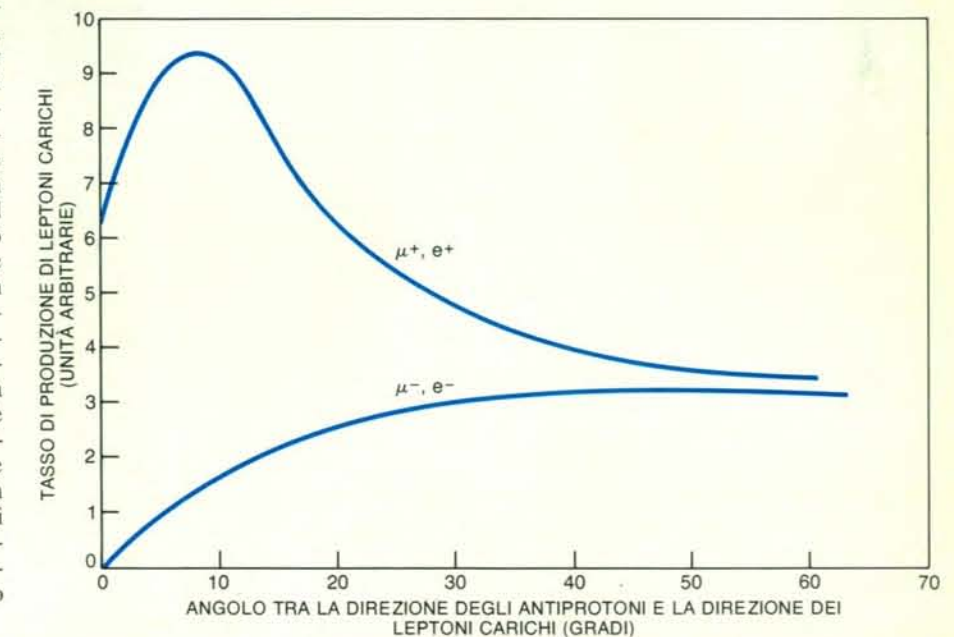
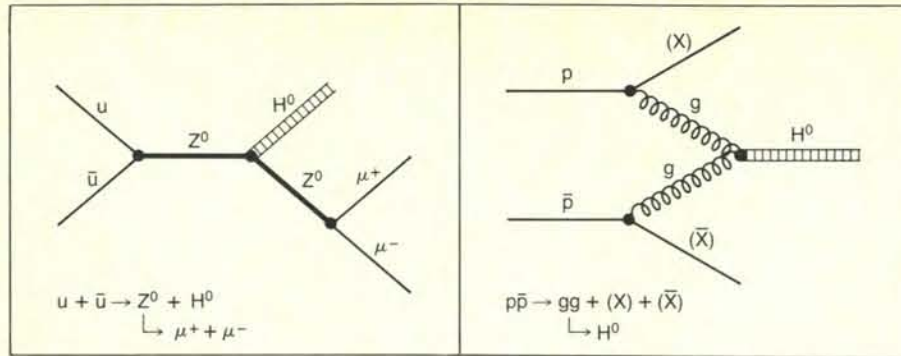
Una possibile comprensione della teoria unificata elettrodebole ha inizio con un immaginario stato primordiale nel quale il fotone e i bosoni vettori intermedi erano tutti ugualmente privi di massa. È stata la rottura di una simmetria della natura a dotare di grandi masse le W^+, le W^- e le Z^0 lasciando il fotone privo di massa. Un meccanismo capace di distinguere in questo modo i portatori delle forze venne trattato per la prima volta nel 1964 da Peter Higgs dell'Università di Edimburgo. È curioso il fatto che il meccanismo sia in grado di fornire massa alle particelle W e Z soltanto postulando un'altra particella pesante, che

L'accumulatore di antiprotoni realizza due funzioni essenziali per gli esperimenti con fasci in collisione del CERN: «ammucchia» i fiotti di antiprotoni iniettati successivamente e li «raffredda» con un processo statistico detto raffreddamento stocastico. Per il raffreddamento l'anello comprende una serie di dispositivi accoppiati, «raccoglitore» e «correttore di rotta» *(forme in grigio)* e i magneti per la curvatura e la focalizzazione dei fasci *(forme bianche)*, presenti in qualsiasi anello di accumulazione. Nel raffreddamento un raccoglitore in una sezione dell'anello di accumulazione capta la deviazione media delle particelle dall'orbita ideale; si invia poi un segnale di correzione attraverso l'anello a un correttore di rotta sul lato opposto, che arriva appena in tempo per riportare le particelle sull'orbita ideale. In figura è indicato uno di questi collegamenti.

L'accumulo e il raffreddamento di antiprotoni nell'anello AA sono illustrati in questa serie di diagrammi, che rappresentano una sezione trasversale dell'anello. Dapprima viene iniettato nell'anello un fiotto di circa 20 milioni di antiprotoni, che è fatto circolare esternamente alla camera a vuoto a larga apertura *(1)*. Durante l'iniezione questo spazio è schermato dal resto della camera da una serranda metallica comandata meccanicamente. Le particelle iniettate vengono preraffreddate con il metodo stocastico per due secondi, riducendo il loro moto casuale di un fattore 10 sia in direzione longitudinale sia in direzione trasversale *(2)*. La serranda viene quindi tolta e gli antiprotoni preraffreddati vengono spostati magneticamente nella posizione di accumulo nel corpo principale della camera *(3)*. La serranda viene nuovamente rimessa e, 2,4 secondi dopo il primo, viene iniettato un secondo fiotto di antiprotoni *(4)*. Il secondo fiotto viene sottoposto allo stesso trattamento, finendo nel «mucchio» dopo esser stato preraffreddato *(5)*. Circa un'ora più tardi, quando sono stati iniettati circa 1500 fiotti e sono stati preraffreddati e ammucchiati circa 30 miliardi di antiprotoni, nel mucchio comincia a formarsi un centro denso *(6)*. Dopo 40 ore, quando sono stati iniettati circa 60 000 fiotti, nel mucchio orbita circa un bilione di antiprotoni, la maggior parte dei quali concentrati al centro *(7)*. Per estrarre questa zona centrale vengono attivati i campi magnetici, fornendo circa 600 miliardi di antiprotoni per gli esperimenti con fasci collidenti. Nell'anello AA rimangono ammucchiati circa 400 miliardi di antiprotoni, che danno inizio al centro successivo *(8)*. Dopo altre 24 ore sarà pronto per l'iniezione il secondo centro di 600 miliardi di antiprotoni. In tutti i casi l'interno dell'anello è sulla destra.

È prevista la comparsa di nuove particelle tra i frammenti di collisioni protone-antiprotone ad alta energia quando un quark della struttura del protone interagisce con un antiquark della struttura dell'antiprotone. Come indica lo schizzo in alto, si presume che il protone sia formato da due quark «su» (chiamati *u*) e da un quark «giù» (*d*), e che l'antiprotone sia formato da due antiquark «anti-su» ($\bar{u}$) e da un antiquark «anti-giù» ($\bar{d}$). Le interazioni quark-antiquark che potrebbero contribuire alla produzione di bosoni vettori intermedi sono mostrate in simboli sulla sinistra e graficamente sulla destra. È postulata l'esistenza di tre bosoni vettori intermedi: due particelle cariche (W^+ e W^-) e una elettricamente neutra (Z^0). Si prevede che tutte e tre abbiano vita media estremamente breve e che siano quindi rivelabili solo mediante i loro prodotti di decadimento, indicati qui genericamente con il simbolo *X*. La prima riga mostra i tre processi più probabili. Le altre quattro righe mostrano processi nei quali vengono creati bosoni vettori intermedi insieme con fotoni energetici (γ). Le ultime due righe mostrano processi in cui si formano coppie di bosoni vettori intermedi. (Quando le particelle sono indicate in moto in entrambe le direzioni lungo una linea verticale l'interazione può procedere secondo entrambi i meccanismi.) I numeri tra parentesi sotto i simboli sulla sinistra danno la carica elettrica di ogni particella; le equazioni mostrano che la carica si conserva in ogni caso. I quark hanno carica frazionaria. I fotoni non ne hanno.

è stata chiamata bosone di Higgs. Anche essa è oggetto di ricerche insieme ai bosoni vettori intermedi.

La teoria elettrodebole ha avuto un'importante conferma sperimentale nel 1973 da una scoperta eseguita al CERN e al Fermi National Accelerator Laboratory (Fermilab) nei pressi di Chicago. Tutte le interazioni deboli della materia allora note comportavano uno scambio di carica elettrica. Per esempio, un protone potrebbe cedere la propria carica +1 a un neutrino (una particella senza massa e senza carica). Il risultato sarebbe che il protone si trasforma in un neutrone e il neutrino si trasforma in un positone, o antielettrone. Tutti questi eventi si possono interpretare come dovuti allo scambio dei bosoni vettori intermedi W^+ e W^-. Nel 1973 gli esperimenti rivelarono interazioni deboli nelle quali le particelle conservano le stesse cariche che avevano prima dell'evento, come accade nelle interazioni elettromagnetiche. Un'interazione debole di questo tipo si può spiegare soltanto con lo scambio di un bosone vettore intermedio neutro (la particella Z^0) o, in una descrizione equivalente, con l'azione di una corrente debole neutra (si veda al riguardo l'articolo *La rivelazione di correnti deboli neutre* di David B. Cline, Alfred K. Mann e Carlo Rubbia in «Le Scienze», n. 80, aprile 1975). Nel 1979 a Weinberg, Salam e Glashow fu conferito il premio Nobel per la fisica «per i loro contributi alla teoria delle interazioni unificate deboli ed elettromagnetiche tra particelle elementari, comprendente... la previsione della corrente debole neutra».

Una volta confermata pienamente l'esistenza delle correnti deboli neutre, diventava naturale cercare un modo per rivelare la Z^0, la W^+ e la W^-. Tuttavia, la creazione di particelle dotate di una massa così grande era un compito difficilissimo. I più grandi acceleratori di particelle dell'epoca erano macchine nelle quali un singolo fascio di protoni viene portato ad alta energia e successivamente indirizzato su un bersaglio fisso. Nella conseguente collisione di una particella del fascio con una del bersaglio la maggior parte dell'energia liberata viene impiegata per muovere il sistema delle due particelle piuttosto che per distruggerlo; soltanto una piccola frazione dell'energia del fascio è disponibile per la creazione di nuove particelle. La sola probabilità di osservare un bosone vettore intermedio, pareva chiaro, era quella di una macchina a fasci collidenti, nella quale le particelle accelerate si scontrano frontalmente, trasformando praticamente tutta la loro energia in nuove particelle.

Gli anelli di accumulazione per elettroni e positoni erano già in funzione da parecchi anni. Il grande vantaggio dell'impiego di elettroni e positoni è che un solo anello di magneti e di cavità a radiofrequenza può accelerare contemporaneamente una particella e la sua antiparticella in direzioni opposte, in modo da formare fasci rotanti in senso opposto in una

La direzione dello spin delle particelle coinvolte nella produzione e nel decadimento dei bosoni vettori intermedi produce un importante effetto sulla direzione di moto dei prodotti di decadimento. Il risultato è un'asimmetria fra le diverse direzioni rispetto ai fasci incidenti, relativamente ai tassi di leptoni carichi osservati. (I leptoni sono particelle, come gli elettroni e i muoni, che prendono parte alle interazioni deboli, ma non a quelle forti.) Per esempio, quando un bosone vettore intermedio carico (W^+ o W^-) decade per formare un muone (μ^+ o μ^-) e un neutrino muonico ($\bar{\nu}_\mu$ o ν_μ), la maggior parte dei muoni carichi positivamente si allontana dalla collisione nella direzione del fascio di protoni incidente (qui, rispettivamente a sinistra e a destra). L'osservazione di questo effetto, che è il solo per interazioni mediate dalla forza debole, verrebbe interpretata come una evidente prova della fugace esistenza di bosoni carichi tra i prodotti di decadimento.

Si prevede un'asimmetria nella distribuzione angolare dei leptoni carichi risultanti da collisioni protone-antiprotone nelle quali vengono creati bosoni vettori intermedi carichi. Qui è rappresentato il tasso di produzione teorico dei leptoni carichi a vari angoli rispetto alla direzione del fascio di antiprotoni. L'effetto è massimo nella direzione in avanti (cioè, a piccoli angoli rispetto alla direzione dell'antiprotone). Si prevede che a questo effetto contribuiscano sia gli elettroni (e^-) e i positoni (e^+), sia i muoni, anche se elettroni e positoni interagiscono con i materiali del rivelatore più rapidamente dei muoni e quindi vengono rivelati con minore frequenza. I punti delle due curve sono stati calcolati per le collisioni protone-antiprotone a un'energia totale di 2000 GeV, l'obiettivo della macchina a fasci collidenti oggi in costruzione al Fermi National Accelerator Laboratory.

Il bosone di Higgs potrebbe fare la sua prima comparsa negli esperimenti con fasci collidenti del CERN o del Fermilab. La scoperta di questa particella pesante e priva di carica (H^0) è considerata la prova definitiva della teoria unificata «standard» che collega le interazioni elettromagnetiche e le interazioni deboli. I diagrammi illustrano due processi che potrebbero portare alla produzione dei bosoni di Higgs. A sinistra, vengono prodotti un bosone di Higgs e un bosone vettore intermedio neutro. A destra, un bosone di Higgs nasce dalla fusione di due gluoni emessi durante una collisione radente tra un protone e un antiprotone. (I gluoni sono le particelle intermedie della forza forte che si pensa tengano assieme i quark all'interno delle particelle del nucleo.)

sola camera a vuoto a forma di ciambella. D'altra parte, dato che gli elettroni e i positoni sono molto leggeri, essi dissipano rapidamente la loro energia quando sono costretti a seguire la traiettoria curva dell'anello di accumulazione. A quel tempo non sembrava possibile la costruzione di una macchina per elettroni e positoni abbastanza grande da raggiungere l'energia dei bosoni vettori intermedi. Il progetto era invece quello di costruire anelli di accumulazione in cui far scontrare frontalmente protoni con altri protoni; per realizzare tali collisioni sono necessari due anelli incrociati. La prima di tali macchine protone-protone, secondo i programmi, doveva entrare in funzione non prima della metà degli anni ottanta o ancora più tardi.

Nel 1976 due di noi (Cline e Rubbia), insieme con Peter M. McIntyre, ebbero un'idea alternativa. Invece di costruire una macchina interamente nuova a fasci collidenti la nostra proposta era quella di realizzare (tra l'altro spendendo molto meno) la conversione di un acceleratore di protoni a bersaglio fisso in una macchina a fasci collidenti, sistemando le cose in modo da produrre un fascio di antiprotoni rotante in senso opposto nello stesso spazio anulare occupato dal fascio di protoni originale. La nostra proposta fu accettata e, dopo una panoramica completa dei problemi in cui il progetto avrebbe potuto incorrere, si decise di costruire macchine protone-antiprotone in due fra i più grandi acceleratori di protoni del mondo: il Superprotosincrotrone (SPS) del CERN, che ha cominciato a funzionare a un'energia di picco di 400 GeV nel 1976, e una versione più progredita di una macchina analoga al Fermilab che all'epoca si trovava ancora in fase di progetto.

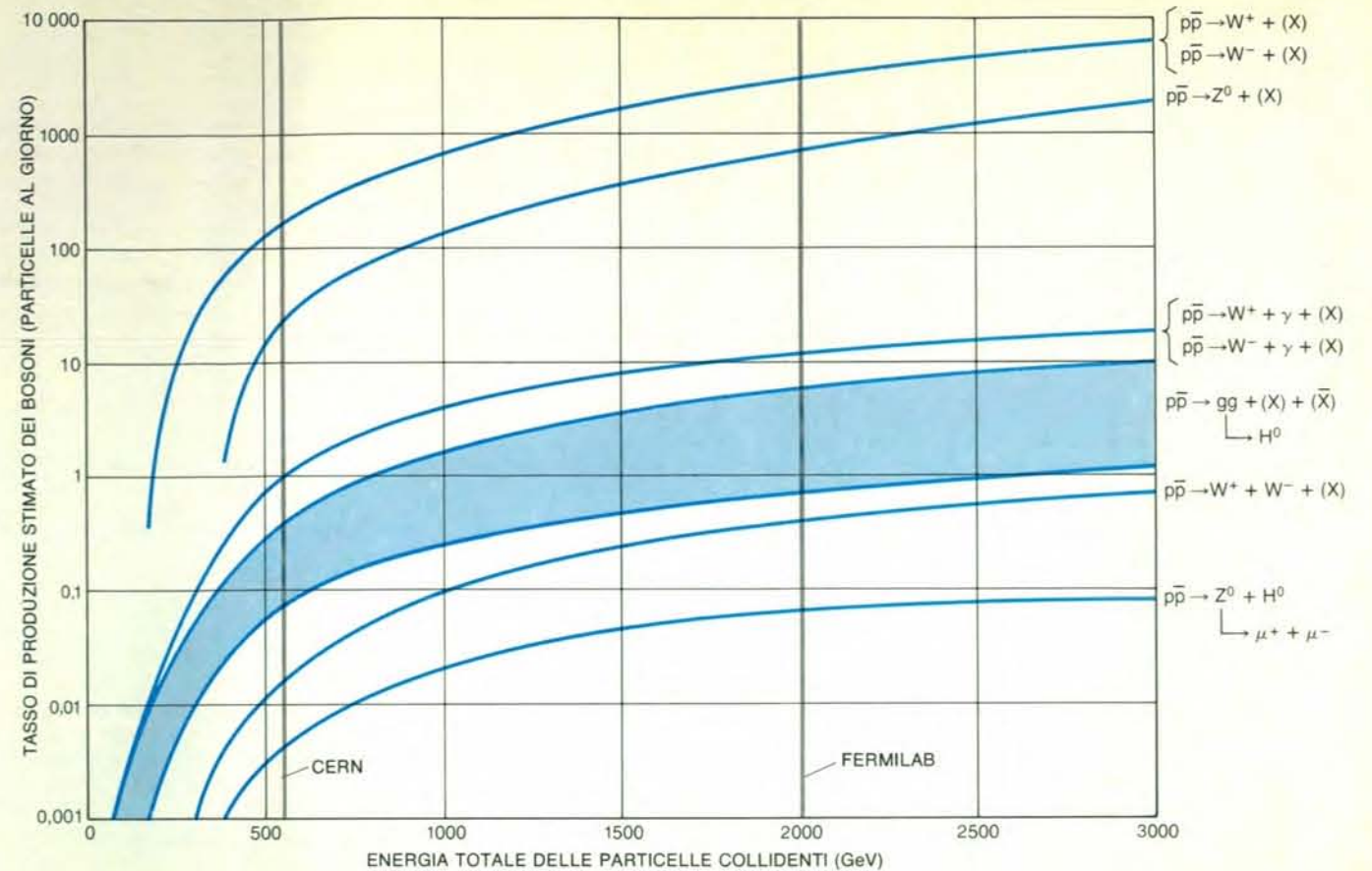
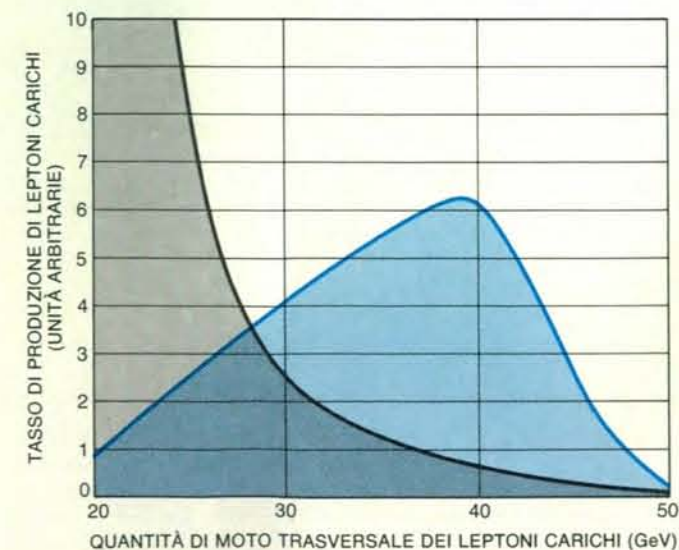
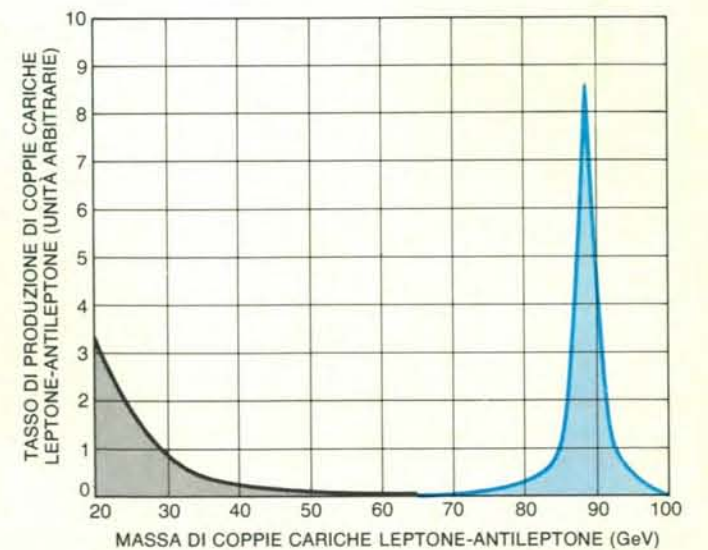
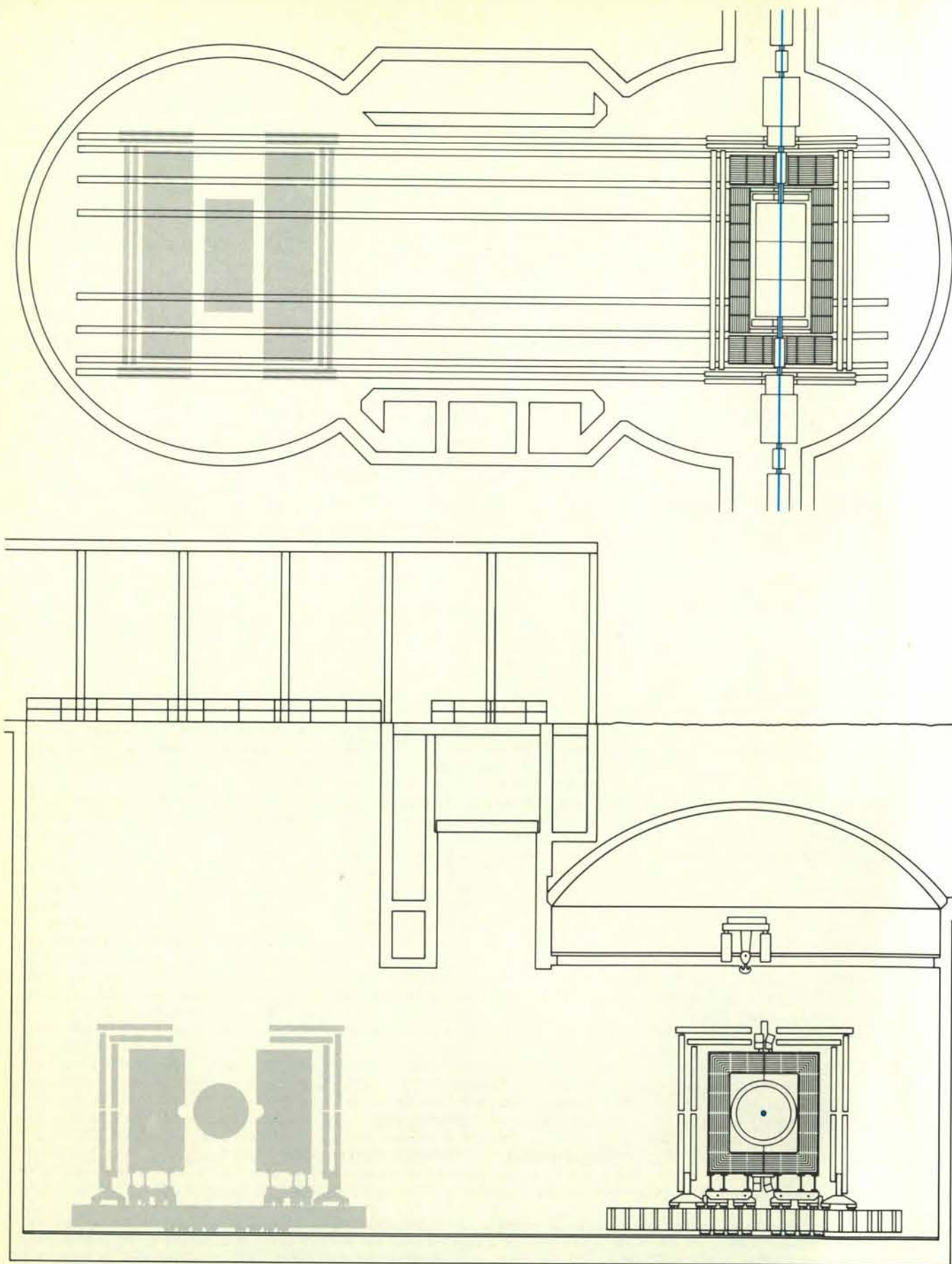
La trasformazione dello SPS fu per molti aspetti più facile da realizzare e fu completata nell'estate del 1981, sotto la direzione di Roy Billinge e di uno di noi (van der Meer); le prime collisioni protone-antiprotone all'energia di picco prefissata di 270 GeV per fascio furono osservate nel mese di luglio. Verso la fine dell'anno, quando la macchina fu fermata per le vacanze di Natale, erano già state registrate più di 250 000 di tali collisioni; a causa della frequenza relativamente bassa alla quale si prevede vengano prodotti i bosoni vettori intermedi in collisioni protoni-antiprotoni, tuttavia, non sorprese il fatto che non ne fosse stato rivelato nemmeno uno in queste prime tornate. Ci aspettiamo però che la situazione si modifichi drasticamente nel corso del prossimo ciclo di esperimenti, cioè quando l'intensità di particelle dei fasci, e pertanto la frequenza delle collisioni, aumenteranno di almeno un ordine di grandezza.

La grande macchina a fasci collidenti protone-antiprotone del Fermilab è ancora in fase di costruzione e si prevede che potrà cominciare a funzionare nel 1985. Poiché il progetto iniziale contemplava l'accelerazione di un solo fascio di protoni a un'energia di 1 TeV (un bilione di elettronvolt), la macchina era stata battezzata Tevatron. Nella sua reincarnazione come macchina a fasci collidenti ci si aspetta che sia in grado di produrre collisioni con un'energia totale di 2 TeV (2000 GeV), contro i 540 GeV della macchina del CERN. La macchina del Fermilab, una volta ultimata, sarà anche il primo grande acceleratore che impieghi un anello di magneti a superconduttori.

Come si fa a produrre fasci opposti di materia e di antimateria in un anello di accumulazione? La parte più difficile, nelle due macchine attuali, consiste nell'accumulare un «fiotto» di antiprotoni sufficientemente denso da assicurare un gran numero di collisioni con i protoni rotanti in senso opposto. Diversamente dai protoni, non è facile ottenere antiprotoni da una sorgente naturale: bisogna produrli a loro volta in collisioni ad alta energia. Si indirizza un fascio di protoni ad alta energia su un bersaglio metallico e gli antiprotoni prodotti nelle collisioni con gli atomi del bersaglio vengono guidati magneticamente in un anello di accumulazione progettato allo scopo. Il processo è estremamente inefficiente; in media viene prodotto un antiprotone di energia relativamente bassa per ogni milione circa di protoni di alta energia incidenti sul bersaglio. Per dare un'idea di tale ritmo di produzione si è calcolato che per poter ottenere un numero utile di collisioni protone-antiprotone nella macchina a fasci collidenti del CERN, si devono raccogliere fiotti di antiprotoni (e di protoni) contenenti ciascuno almeno 100 miliardi di particelle. I successivi fiotti di antiprotoni vengono raccolti e «ammucchiati» ogni 2,4 secondi; con questo ritmo ci vogliono circa 24 ore per poter accumulare alcune centinaia di miliardi di antiprotoni per gli esperimenti a fasci collidenti del CERN.

La creazione di un numero sufficiente di antiparticelle non è il solo problema. Gli antiprotoni emergono dal bersaglio con velocità diversa e puntando in direzioni diverse. Osservati nel loro stesso sistema di riferimento gli antiprotoni formano un gas, e i loro moti casuali costituiscono un indice della temperatura. Se la temperatura è troppo alta, alcune particelle colpiranno le pareti dell'acceleratore e il fascio sarà distrutto. È necessario quindi ricorrere a qualche metodo per «raffreddare» il fascio di antiprotoni (per ridurne, cioè, i moti casuali) in modo da mantenerlo il più concentrato possibile prima del suo ingresso nell'acceleratore.

Una tecnica di raffreddamento del fascio, chiamata raffreddamento elettronico, venne proposta per la prima volta più di dieci anni fa da Gersh I. Budker dell'Istituto di fisica nucleare di Novosibirsk in Unione Sovietica. Essa consiste in pratica nel mescolare un fascio «freddo» di elettroni (cioè un fascio nel quale tutte le particelle hanno la stessa velocità e la stessa direzione) con il fascio di antiprotoni «caldo» per un breve tratto. Nel corso di tale processo una parte dell'energia termica casuale degli antiprotoni viene trasferita agli elettroni. Il ripetuto mescolamento del fascio di antiprotoni con nuovi fasci di elettroni può raffreddare significativamente gli antiprotoni, purché la loro energia iniziale non sia troppo alta. In pratica accade che lo schema del CERN richiede un fascio di antiprotoni la cui energia è inizialmente troppo alta per poterlo raffreddare efficacemente con questo metodo e non viene più studiato per questo scopo. Il raffreddamento elettronico continua tuttavia ad essere argomento di studio per il suo possibile impiego nel progetto del Fermilab.

Un altro metodo di raffreddamento del fascio, più adeguato alle esigenze della macchina protone-antiprotone del CERN, venne inventato nel 1968 da uno di noi (van der Meer). Questo meto-

I tassi di produzione stimati per bosoni vettori intermedi e per bosoni di Higgs in vari processi di collisione variano in funzione dell'energia totale dei fasci collidenti. Le curve indicano il tasso calcolato per ciascun processo per un giorno di funzionamento al tasso di collisione previsto nel progetto delle macchine a fasci collidenti del CERN e del Fermilab. (Il tasso effettivamente rivelato sarà inferiore a causa dei vari effetti sperimentali di fondo.) Le due rette verticali in grigio indicano l'energia totale prevista per le collisioni protone-antiprotone al CERN e al Fermilab. I processi risultanti nella produzione di un bosone vettore intermedio carico accompagnato da un fotone energetico si prevede abbiano la funzione di indicatori sensibili alle proprietà magnetiche delle particelle W^+ e W^-. Il processo in cui viene prodotto un bosone di Higgs per fusione di due gluoni è meno prevedibile degli altri processi e quindi il tasso relativo viene rappresentato da un intervallo di valori.

Ci si aspetta che i segnali indicatori della produzione di bosoni vettori intermedi *(in colore)* risaltino sul «rumore» di fondo *(in grigio)*, in particolare a grandi angoli rispetto all'asse del fascio. Il grafico sulla sinistra mostra lo spettro di massa calcolato dei leptoni carichi che dovrebbe essere emesso con una grande quantità di moto trasversale da collisioni protone-antiprotone in cui si creano bosoni vettori intermedi. Si prevede che il picco appaia circa in corrispondenza della metà della massa stimata di un bosone vettore intermedio. Secondo le previsioni il tasso di fondo dei leptoni dovuti ad altre sorgenti correlate al processo di collisione non presenta picchi ed è inferiore al tasso previsto per il decadimento delle particelle W^+ e W^-. Il grafico sulla destra mostra lo spettro di massa calcolato per il decadimento di un bosone vettore intermedio neutro in una coppia di leptoni carichi. In questo caso lo spettro presenta un picco nei pressi della massa prevista per la particella Z^0 (90 GeV). Qui il rumore di fondo è trascurabile, pertanto questo processo è l'unico nel quale è molto probabile scoprire la particella Z^0.

Il rivelatore UA1 è installato in un'area sperimentale a circa 25 metri di profondità. Sia nella veduta in pianta *(in alto)* sia in quella laterale *(in basso)* il rivelatore è disegnato a tratti neri quando si trova nella sua posizione di funzionamento a cavallo del tubo del fascio dell'SPS e a blocchi in grigio quando è parcheggiato nel suo garage. L'UA1 è di gran lunga il più grande rivelatore finora costruito per un esperimento a fasci collidenti: è lungo 10 metri e largo cinque e pesa 2000 tonnellate. Il costo di costruzione è stato di circa 20 milioni di dollari.

do, chiamato raffreddamento stocastico (perché si basa su un processo statistico), impiega un dispositivo di «raccolta» o sensore in una sezione dell'anello di accumulazione per misurare la deviazione media delle particelle dall'orbita ideale. La misura viene convertita in un segnale di correzione, il quale viene trasmesso attraverso l'anello a un dispositivo «correttore di rotta» sul lato opposto. Il correttore di rotta applica un campo elettrico alla propria sezione dell'anello in tempo per riportare verso l'orbita ideale il centro di massa delle particelle che passano. Anche se le particelle si muovono con una velocità quasi pari a quella della luce, il segnale di correzione può arrivare in tempo perché esso percorre il cammino più breve lungo una corda dell'anello di raffreddamento.

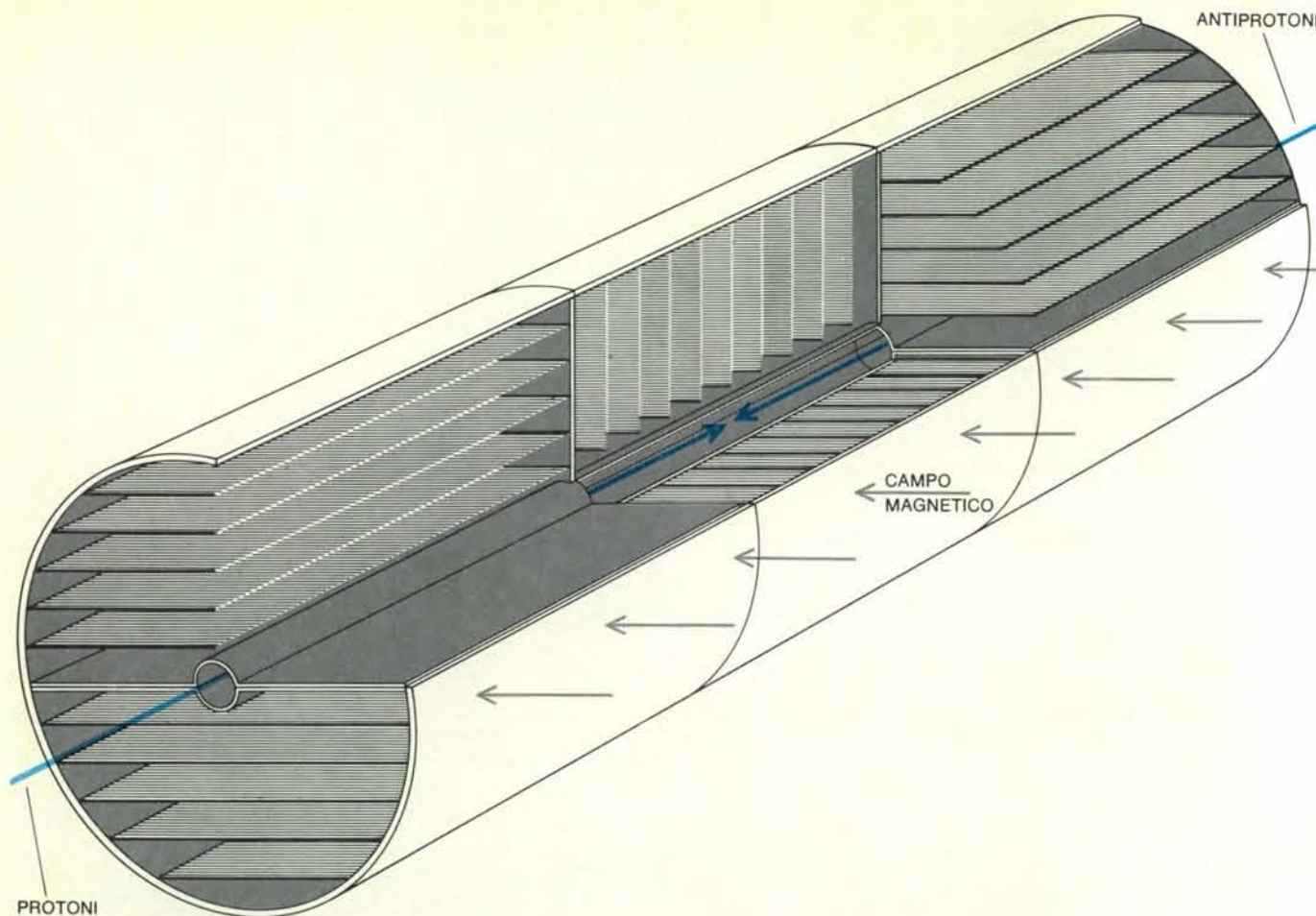
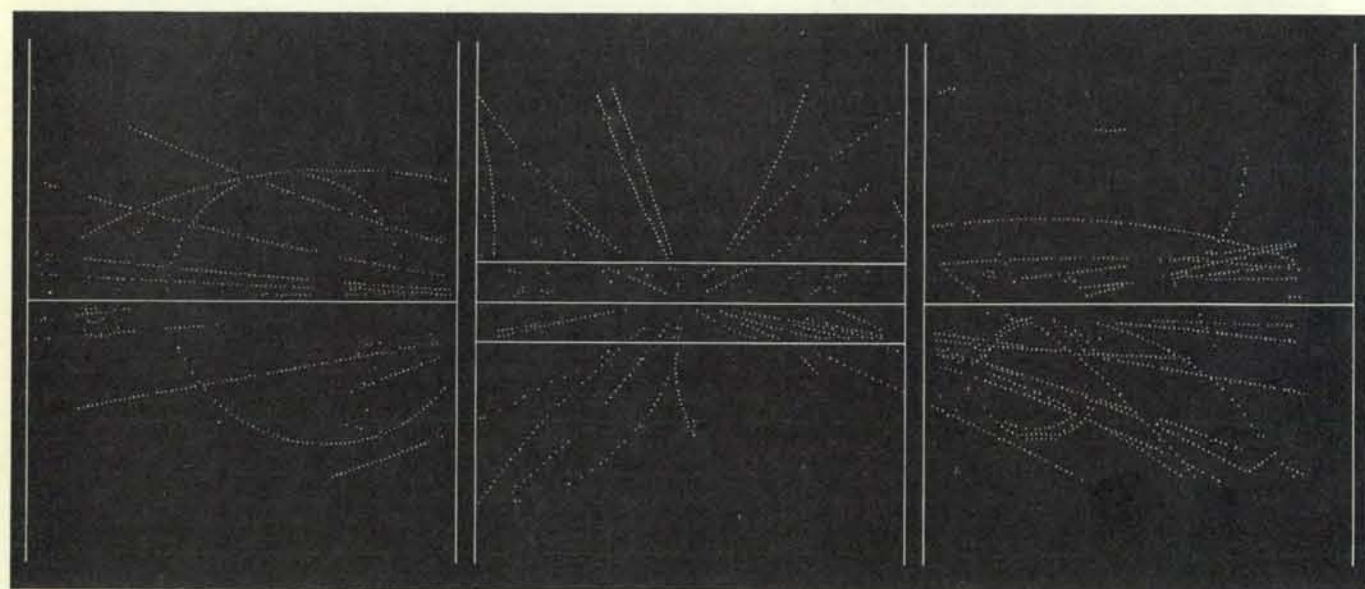
Entrambe le tecniche di raffreddamento del fascio sono state sperimentate con successo negli ultimi anni a Novosibirsk, al CERN e al Fermilab. Abbiamo quindi tutti i motivi per credere che gli anelli collettori di antiprotoni a grande scala, come quelli del CERN e del Fermilab, funzioneranno come da progetto. Il raffreddamento dei fasci sta diventando una fase normale della tecnologia degli acceleratori.

Negli esperimenti del CERN le particelle vengono guidate attraverso una complessa serie di dispositivi per la manipolazione del fascio che sono tra loro collegati. Prima un fascio di protoni viene accelerato a un'energia di 26 GeV nel Protosincrotrone (PS), l'originale anello di accelerazione del CERN ultimato nel 1959. Il fascio di protoni viene quindi indirizzato su un bersaglio di rame, producendovi un getto di particelle, tra le quali un piccolo numero di antiprotoni con una energia di 3,5 GeV. Gli antiprotoni vengono raccolti e trasferiti a un anello di accumulazione ad ampia apertura chiamato Accumulatore di antiprotoni (AA), nel quale essi vengono dapprima preraffreddati con il metodo stocastico e quindi spostati su un'orbita leggermente più piccola, dove vengono aggregati ai fiotti precedentemente iniettati e sottoposti poi a ulteriore raffreddamento. Una volta raccolti alcuni miliardi di antiprotoni, essi vengono rinviati all'anello PS, dove vengono accelerati a 26 GeV prima di essere iniettati nello SPS. Nel frattempo i protoni da 26 GeV dell'anello PS vengono iniettati nell'anello SPS in direzione opposta. Infine, i due fasci rotanti in senso opposto vengono accelerati a 270 GeV nell'anello SPS. I fasci si scontrano in due luoghi di interazione, dove sono sistemati i grandi rivelatori di particelle *(si vedano le illustrazioni a pagina 25)*. Le interazioni sono così rare che la vita media del fascio, che è di parecchie ore, non ne viene modificata.

Una vista laterale del rivelatore UA1 lo mostra in posizione sulla linea del fascio dell'SPS. Sono indicate le componenti sperimentali per vari obiettivi, fra cui la ricerca dei bosoni vettori intermedi.

Attualmente il progetto del Fermilab richiede una sorgente di antiprotoni più intensa di quella in esercizio al CERN. Per la produzione di antiprotoni al Fermilab si impiegheranno protoni di energia più elevata, mentre si stanno studiando numerosi schemi alternativi. Uno di essi si basa su una combinazione di raffreddamento stocastico e di raffreddamento elettronico per accumulare 100 miliardi di antiprotoni in meno di un'ora. Gli antiprotoni dovrebbero essere iniettati nell'anello del Tevatron, accelerati a un'energia di 1000 GeV e fatti entrare in collisione, nelle due aree sperimentali, con i protoni da 1000 GeV rotanti in senso opposto. La prima area, la cui costruzione è previsto debba iniziare subito, è progettata per contenere un grandissimo rivelatore.

Secondo la teoria elettrodebole, i bosoni vettori intermedi possono essere creati in collisioni protone-antiprotone attraverso vari meccanismi in cui un quark della struttura del protone interagisce con un antiquark della struttura dell'antiprotone. Sia il protone sia l'antiprotone si ritengono costituiti da tre particelle; nella curiosa nomenclatura della teoria QCD si dice che il protone comprende due quark «su» (indicati con u) e un quark «giù» *(d)*, mentre l'antiprotone possiede due antiquark «anti-su» ($\bar{u}$) e un antiquark «anti-giù» ($\bar{d}$). Quando un quark e un antiquark entrano in collisione, si annichilano a vicenda creando un fiotto di energia che può rimaterializzarsi sotto forma di nuove particelle, tra le quali bosoni vettori intermedi. In alcuni casi è previsto che debba comparire un solo bosone vettore intermedio (accompagnato da altri tipi di particelle); in altri casi è prevista la comparsa di una coppia di bosoni vettori intermedi *(si veda l'illustrazione a pagina 26)*.

Il tasso di produzione previsto per le particelle W^+, W^- e Z^0 nelle collisioni protone-antiprotone varia in funzione di un parametro sperimentale chiamato luminosità, definito come il numero di particelle di alta energia che attraversano nell'unità di tempo la sezione trasversale della regione di interazione (numero di particelle per centimetro quadrato per secondo). In base al progetto, supponendo di iniettare 600 miliardi di antiprotoni per fiotto, la luminosità della macchina del CERN dovrebbe essere di 10^{30} particelle per centimetro quadrato per secondo. Con lo stesso numero di antiprotoni per fiotto, la macchina del Fermilab dovrebbe invece raggiungere una luminosità di 4×10^{30} particelle per centimetro quadrato per secondo (a causa della sua più elevata energia e quindi delle minori dimensioni del fascio). Si può calcolare che a tali luminosità il tasso di produzione di particelle W^+, W^- e Z^0, sia singole che a coppie, dovrebbe essere sufficientemente alto da farle rivelare molto spesso, forse addirittura migliaia di volte al giorno *(si veda il grafico a pagina 29 in alto)*. Poco prima di essere fermata in dicembre la macchina del CERN ha raggiunto per poco tempo una luminosità di 10^{28} particelle per centimetro quadrato

Il sistema di rivelazione centrale dell'UA1 è costituito da «tre camere a drift» (a deriva) cilindriche, ciascuna delle quali contiene una fitta schiera di fili e un gas a bassa pressione. In tutte e tre le camere i fili sono tesi orizzontalmente. Nella camera centrale i fili orizzontali sono disposti in piani verticali; nelle due camere attigue essi sono disposti in piani orizzontali. Una particella carica che attraversa la camera ionizza le molecole del gas, il quale tende poi a migrare verso i fili, depositandovi la carica. La distribuzione delle cariche che appaiono su molti fili viene registrata elettricamente e può successivamente essere analizzata da un calcolatore per ricostruire la traiettoria della particella sullo schermo di un tubo a raggi catodici. Il diametro delle camere è all'incirca di tre metri. I fili distano circa tre millimetri uno dall'altro.

Questa immagine è la registrazione visiva di una collisione protone--antiprotone avvenuta alla fine del 1981 nel sistema centrale di rivelazione dell'UA1 ed è stata eseguita fotografando una visualizzazione prodotta mediante un calcolatore. L'evento è lo stesso che appare sulla copertina di questo fascicolo. Finora il calcolatore ha registrato oltre 250 000 eventi di questo tipo. Ne verranno registrati altri milioni, con la ripresa della ricerca di bosoni vettori intermedi al CERN. Il campo magnetico applicato fa incurvare le traiettorie delle particelle cariche.

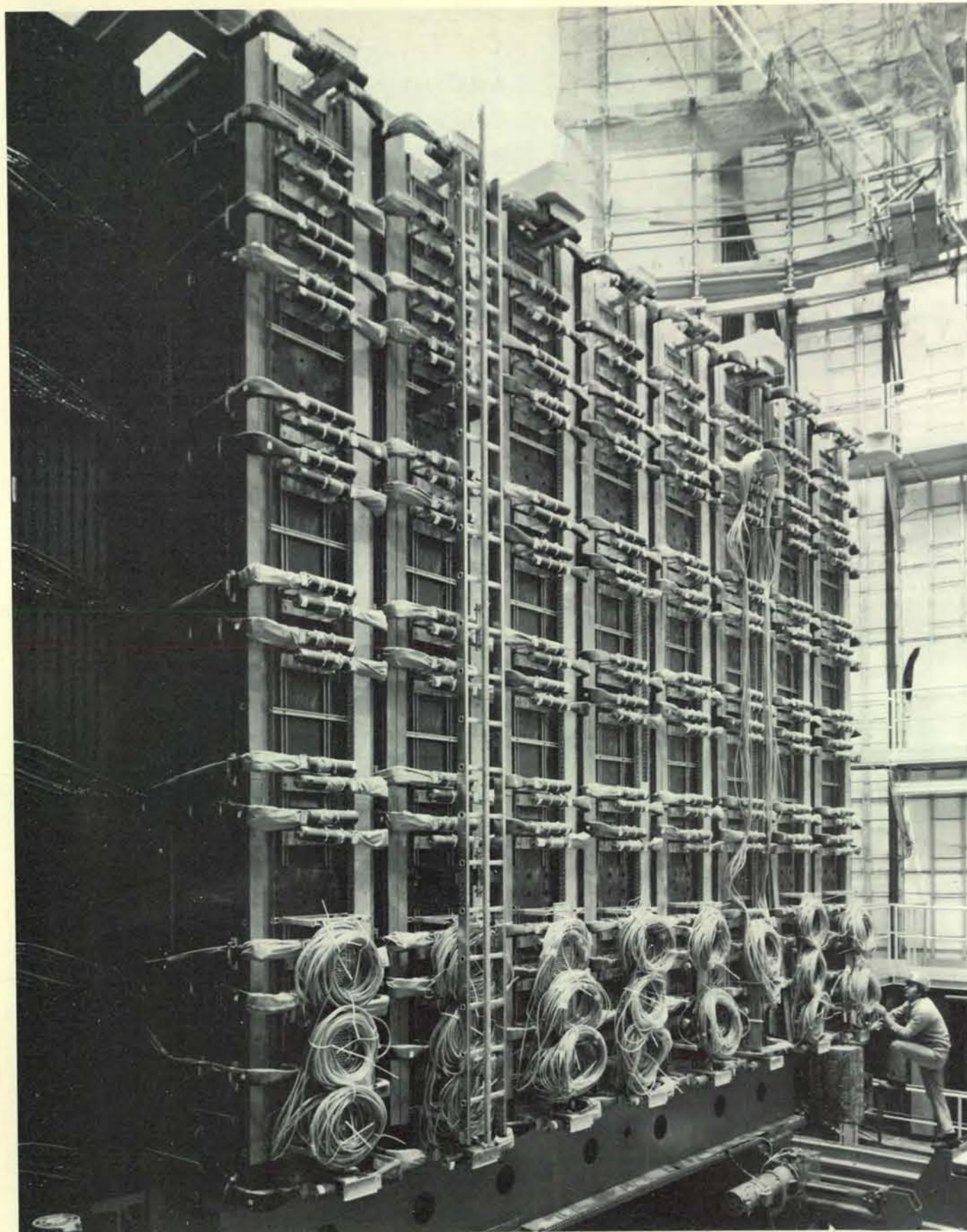
per secondo, arrivando così alla soglia delle condizioni in cui ci si aspetta di vedere per la prima volta di sfuggita un bosone vettore intermedio. (Anche le prime esperienze di funzionamento della macchina a fasci collidenti del CERN hanno mostrato che le collisioni protone--antiprotone si possono ottenere in linea di principio a luminosità di almeno 10^{31} particelle per centimetro quadrato per secondo, purché sia disponibile un numero di antiprotoni sufficiente.)

In che modo i bosoni vettori intermedi prodotti in tali collisioni riveleranno la loro presenza? Si prevede che la vita media di tali particelle debba essere estremamente breve. In circa 10^{-20} secondi esse dovrebbero decadere per formare una varietà di altre particelle, principalmente coppie quark-antiquark e coppie leptone-antileptone. (I leptoni sono particelle sensibili alla forza nucleare debole, ma non a quella forte.) I leptoni carichi, per esempio gli elettroni e i muoni, si possono rivelare in molti modi. In generale l'obiettivo consiste nel rivelare i leptoni carichi dovuti al decadimento dei bosoni vettori intermedi e nel confrontare il loro tasso di produzione o altre loro proprietà con i valori previsti per leptoni dovuti ad altre fonti, sempre generate nel processo di collisione. Si prevede che il segnale, in questo caso la distribuzione angolare dei leptoni del decadimento di particelle W^+, W^- e Z^0, risalti sul «rumore» di fondo di leptoni provenienti da altre sorgenti, particolarmente a grandi angoli rispetto all'asse del fascio *(si veda l'illustrazione a pagina 29 in basso).*

Un'indicazione inconfondibile della presenza di bosoni vettori intermedi dovrebbe essere la comparsa di una marcata asimmetria nell'intensità di conteggio dei leptoni nelle direzioni «in avanti» e «all'indietro» (misurate, arbitrariamente, rispetto alla direzione del fascio di antiprotoni). I leptoni prodotti direttamente da interazioni forti o elettromagnetiche del fascio di particelle dovrebbero essere del tutto simmetrici. Invece, secondo la teoria elettrodebole, i bosoni vettori intermedi che decadono dovrebbero emettere leptoni a carica positiva prevalentemente in avanti e leptoni a carica negativa prevalentemente all'indietro. La prevista asimmetria dei leptoni, che è la sola per eventi mediati dalla forza debole, è dovuta agli spin delle particelle coinvolte nella produzione e nel decadimento di bosoni vettori intermedi *(si vedano le illustrazioni a pagina 27).* L'osservazione di tale effetto proverebbe chiaramente che finalmente i tanto cercati bosoni vettori intermedi sono stati scoperti. Si potrebbero poi misurare le loro altre proprietà.

La prova definitiva della correttezza della teoria elettrodebole dovrebbe essere l'osservazione del bosone di Higgs tra i frammenti delle collisioni protone-antiprotone. La scoperta di tale particella dimostrerebbe non solo che la forza elettromagnetica e quella debole sono unificate, ma anche che l'unificazione è del tipo previsto dalla teoria elettrodebole standard. Una trattazione approfondita delle tecniche sperimentali necessarie per la rivelazione del bosone di Higgs va oltre lo scopo di questo articolo. Tuttavia, le collisioni protone-antiprotone potrebbero dare origine a bosoni di Higgs e la frequenza di produzione calcolata è sufficientemente alta perché essi possano essere scoperti mediante i loro caratteristici prodotti di decadimento sia nella macchina a fasci collidenti del CERN sia in quella del Fermilab.

Un altro grande rivelatore, l'UA2, è stato recentemente ultimato al CERN. È stato progettato più specificamente per la ricerca di bosoni vettori intermedi. Diversamente dall'UA1, non ha campi magnetici. È installato in una seconda area sperimentale a più di 60 metri di profondità, a una certa distanza dall'UA1. Nella fotografia il rivelatore è posizionato sulla linea del fascio dell'SPS; è dotato di ruote e, quando non è in funzione, viene riportato nella grande area di lavoro in primo piano. Il grande albero cilindrico in alto viene usato per calare componenti pesanti dalla superficie.

Sono stati progettati numerosi rivelatori a grandi dimensioni per la ricerca dei prodotti di decadimento dei bosoni vettori intermedi e dei bosoni di Higgs. Uno di questi dispositivi, chiamato UA1 (*Underground Area 1*, area sotterranea 1), è ultimato e pronto a entrare in esercizio al CERN. Il rivelatore è il risultato di uno sforzo di collaborazione di più di 100 fisici di 11 istituzioni europee e statunitensi: l'Università di Aquisgrana, il Laboratorio di fisica delle particelle di Annecy, l'Università di Birmingham, il CERN, il Queen Mary College (Londra), il Collège de France (Parigi), l'Università della California (Riverside), l'Università di Roma, il Rutherford Laboratory, il Centro di ricerche nucleari di Saclay e l'Università di Vienna. È lungo 10 metri e largo cinque e il suo peso totale è di 2000 tonnellate. La sala sotterranea in cui è installato, a circa 25 metri di profondità, è sufficientemente grande da consentire di trasportarlo su ruote in un «garage» quando non

Un nuovo rivelatore di particelle è stato progettato e costruito al CERN di Ginevra da una squadra di più di 100 fisici di 11 istituzioni europee e statunitensi. Con l'aiuto di questo dispositivo si prevede di poter eseguire la prima osservazione di un bosone vettore intermedio. Il grande e versatile rivelatore, chiamato UA1 *(Underground Area 1)*, appare qui nel «garage» adiacente al più grande acceleratore di particelle del CERN: il Superprotosincrotrone (SPS), recentemente trasformato in una macchina a fasci collidenti protone-antiprotone. Quando il rivelatore è pronto a entrare in funzione, viene spostato verso sinistra, mediante ruote che scorrono su binari, verso la traiettoria dei fasci collidenti. L'attrezzatura elettronica copre la parte esterna dell'apparecchiatura, nascondendo le camere di rivelazione centrali.

deve essere posto sulla traiettoria dei fasci in collisione.

Il rivelatore UA1 è un dispositivo versatile progettato per essere sensibile a molti tipi di particelle e per raccogliere informazioni in un ampio angolo solido attorno al punto in cui i fasci entrano in collisione. Esso misura l'energia delle particelle in molti modi, in particolare in base alla curvatura delle loro traiettorie in un campo magnetico. Un grande dipolo magnetico applica orizzontalmente il campo magnetico principale in un volume di 85 metri cubi.

All'interno del magnete, attorno al tubo del fascio, vi sono le «camere a drift» (a deriva), ciascuna delle quali contiene una serie di fili strettamente spaziati e un gas a bassa pressione. Una particella. La camera a drift centrale ha attraversa la camera ionizza le molecole del gas; gli ioni migrano quindi verso i fili, sui quali depositano la loro carica. Dalla distribuzione delle cariche su molti fili si può ricostruire la traiettoria della particella. La camera a drift centrale ha i fili disposti in piani verticali, mentre le due camere adiacenti hanno i fili disposti in piani orizzontali. I segnali delle particelle che attraversano i piani dei fili possono essere elaborati da un calcolatore per produrre un'immagine dei prodotti di decadimento sullo schermo di un tubo a raggi catodici *(si vedano le illustrazioni a pagina a 32)*.

Attorno alle tre camere a drift vi sono vari altri rivelatori. Appena al di fuori del rivelatore più interno si trova un calorimetro a piombo, un dispositivo che misura la quantità di energia depositata in esso da una particella carica quale un elettrone. Il calorimetro è a sua volta circondato da una serie di lastre di ferro alternate a contatori a scintillazione per misurare l'energia di particelle più pesanti, quali i pioni, mediante le loro interazioni con gli atomi di ferro delle lastre. Infine, all'esterno dell'apparecchiatura vi sono molte grandi camere per rivelare i muoni che attraversano sia il piombo sia le lastre di ferro.

Un altro grande rivelatore, l'UA2, è stato progettato più specificamente per la ricerca dei bosoni vettori intermedi. Esso non ha campo magnetico, ma si basa invece su una grande schiera di calorimetri simili a quelli del rivelatore UA1 per misurare l'energia e la direzione delle particelle emergenti. È installato in una seconda area sperimentale, a una profondità di più di 60 metri *(si veda l'illustrazione a pagina 33)*. Sono in programma al Fermilab altri rivelatori simili all'UA1 e all'UA2. Se i bosoni vettori intermedi esistono, siamo convinti che questi rivelatori siano in grado di scoprirli e di indagare sulle loro proprietà, confermando quindi la teoria elettrodebole unificata. Anche i bosoni di Higgs, se esistono, potrebbero essere rivelati, fornendo pertanto ulteriore sostegno alla teoria. Naturalmente, è anche possibile che la teoria elettrodebole sia sbagliata e che nessuna di tali particelle esista. In un modo o nell'altro, la risposta non dovrebbe farsi attendere molto.

L'oro invisibile

La scoperta di giacimenti con corpuscoli d'oro di dimensioni inferiori a due micrometri e analizzabili solo al microscopio elettronico o con metodi chimici ha iniziato una nuova fase dell'esplorazione aurifera

di Giancarlo Facca

Quando si nomina questo metallo nobile, si evocano immediatamente ricordi romantici, drammi, avventure d'altri tempi e le follie collettive delle corse all'oro in California, Alaska, Australia. Questi episodi hanno avuto un'importanza storica considerevole nell'ultima metà del secolo scorso, non tanto per il rapido apporto di ricchezza, cosa che pure a volte si è verificata, quanto per le migrazioni di massa in paesi praticamente deserti.

La corsa all'oro esemplare, e la più nota, fu quella della California, che cominciò nel 1848 e raggiunse il massimo vigore nei due decenni che seguirono. Dapprima vi parteciparono i pochi americani che vivevano nella regione, in conseguenza delle prime sensazionali scoperte nelle valli del fiume Sacramento e dei suoi affluenti. Per qualche anno, quasi tutte le attività cessarono: i contadini abbandonavano i campi già pronti per il raccolto, i soldati disertavano e perfino i marinai lasciavano le navi da trasporto e da guerra per correre alla Sierra Nevada! Alla fine del 1848, circa diecimila uomini stavano scavando con il piccone e la pala i placer e i filoni quarziferi. Usavano la bateia, una specie di larga scodella dal margine svasato, e con questo primitivo strumento separavano l'oro dalla sabbia o dal quarzo frantumato, mescolando il tutto con acqua. Muovendo opportunamente la bateia a circoli, il materiale sterile esce con l'acqua dai bordi, mentre l'oro, che è più pesante, resta sul fondo.

Durante l'anno seguente, il 1849, la corsa ai campi d'oro si estese e gli Argonauti (come vennero chiamati) si muovevano dalla costa atlantica per raggiungere la costa pacifica attraversando l'immenso continente con i lenti carri, oppure si imbarcavano sulle navi lungo la rotta di Capo Horn o quella dell'istmo dell'America Centrale. Dall'Europa, s'imbarcavano ogni giorno per l'America un migliaio di emigranti; altre navi partivano dalle Hawaii, dall'America Latina, dalla Cina, dalla Russia e perfino dalle isole Marchesi. Giunti in California, questi Argonauti incontravano altre difficoltà prima di raggiungere la regione aurifera, ma la speranza di una rapida ricchezza, unita al senso di euforia connesso con l'abbandono della vecchia vita tediosa e le forti inibizioni delle loro comunità tradizionali, generavano negli emigranti una ribalda allegria. Cantavano:

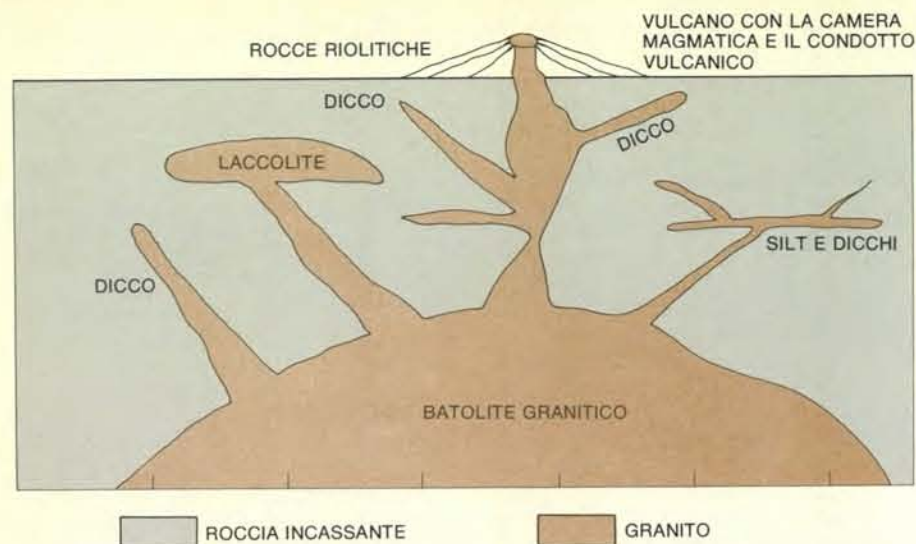
Oh, California,
That's the land for me!
I am going to Sacramento
With a washbowl on my knee.

(Oh, California, questo è il paese per me! Io vado nella valle del fiume Sacramento, con la bateia sulle mie ginocchia).

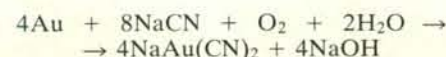
Tanto entusiasmo era giustificato? Almeno per i primi anni, la risposta è affermativa. Prima di iniziare la corsa all'oro, il retaggio di questi uomini era la vita dura dell'agricoltore. Per chi lavorava sotto padrone, due o tre dollari al giorno erano una buona paga, mentre «tra diecimila uomini che lavoravano agli scavi, nell'autunno del 1848, soltanto pochi non riuscivano a racimolare almeno venti dollari al giorno, e centinaia potevano vantare scoperte favolose. Jacob Leese e quattro compagni fecero 75 000 dollari in tre mesi nella valle del fiume Yuba, usando il piccone e la pala. William Daylor scavò 15 000 dollari in una settimana; al Parks Bar il ricavo medio per uomo era di 100 dollari al giorno...» (F. Riesemberg, Jr.); e non mancavano i ritrovamenti spettacolosi di grosse pepite pesanti anche qualche chilogrammo; una famosa pepita, *The Monumental,* aveva un peso di circa 7 chilogrammi; il suo valore, a quel tempo (a 20 dollari all'oncia troy) era di 4500 dollari. Ma questa pepita appare un nano rispetto alle molte, di oltre 70 chilogrammi, trovate a fior di terra in Australia. Per quanto riguarda l'ammontare complessivo dell'oro prodotto in California prima del 1900, si fa la cifra di un miliardo di dollari di quel tempo, calcolando 20 dollari all'oncia. Ai prezzi dei primi mesi del 1982 (circa 300 dollari all'oncia, ovvero 10 dollari al grammo) la stima sale a 15 miliardi di dollari.

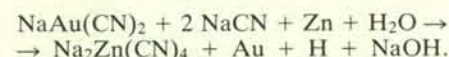
L'oro prodotto dai giacimenti della California e da quelli simili in altri paesi produttori è visibile a occhio nudo e quindi si può separare dalle rocce sterili che lo contengono con procedimenti meccanici abbastanza semplici; ma le cose si complicano quando le pagliuzze d'oro sono troppo piccole per essere visibili a occhio nudo. Bisogna allora ricorrere per l'estrazione a sistemi chimici, il più comune dei quali è il processo di cianurazione, introdotto nel 1887 da Y. S. MacArthur e R. W. e W. Forest, nel quale si usa una soluzione debole (0,01 - 0,10 per cento) di cianuro di potassio equivalente. Il cianuro di potassio è un composto di formula KCN estremamente velenoso anche in piccole quantità; al principio di questo

Cercatori d'oro della California verso la fine del secolo scorso. L'uomo in primo piano usa il *gold pan* (la bateia), che è una specie di scodella svasata: vi si mettono la sabbia o il quarzo triturato e si colma d'acqua, mescolando bene. Con un movimento rotatorio o facendo oscillare lo strumento, si facilita la sedimentazione dell'oro, che è più pesante, al fondo della scodella. Poi, inclinando con cura lo strumento e facendolo oscillare, si eliminano l'acqua e i granelli di roccia, si concentra e si raccoglie l'oro rimasto al fondo. In secondo piano, un pozzo rudimentale e, a destra, il *long tom:* un canale di legno lungo da 2 a 4 metri e largo da 30 a 60 centimetri, munito di traverse o di scanalature. Il minerale è mosso lungo il piano inclinato da una lenta corrente d'acqua e l'insieme viene rimescolato di continuo. L'oro viene trattenuto dalle scanalature del fondo.

Un dispositivo per la concentrazione dell'oro nei giacimenti tipo placer (giacimenti di età in genere recente, di natura piuttosto semplice, che hanno un'origine alluvionale). Un canale porta l'acqua, che muove la grande ruota, da un livello più alto del torrente. Un ingranaggio fa girare il primitivo agitatore della vasca, nella quale un minatore getta ciottoli e sabbia del giacimento, provvedendo continuamente al rimescolamento. Un secondo canale porta l'acqua necessaria per asportare le parti più leggere della torbida, la quale viene a sua volta concentrata nello sdrucciolo munito di scanalature trasversali, che è possibile vedere in primo piano. Il disegno, realizzato da H. C. Smith, è tratto dal volume *Geology of Northern California*, una pubblicazione curata dalla California Division of Mines and Geology.

I giacimenti d'oro sono correlati alle intrusioni granitiche penetrate nei livelli relativamente superficiali della crosta terrestre. Le dimensioni di un batolite granitico, cioè di una grande massa di magma, sono enormi: i batoliti si estendono su migliaia, alcuni addirittura su milioni di chilometri quadrati. Il batolite della Sierra Nevada, che è all'origine dei giacimenti del Mother Lode, è lungo circa 1000 chilometri da nord a sud e largo circa 200 da ovest a est. Si è formato con apporti di magma, che si sono intrusi l'uno nell'altro, in fasi successive, durante circa 100 milioni di anni, dal Giurassico superiore al Cretaceo superiore (da 160 a 60 milioni di anni fa). In una fase geologica successiva, l'erosione ha spianato parte della roccia incassante e parte dello stesso batolite, ormai in superficie *(si veda anche l'illustrazione in alto nella pagina a fronte).*

secolo, però, fu scoperto in Germania un metodo economico per produrre il cianuro di sodio, che ora è generalmente usato. Anche questo composto è molto velenoso e, come il cianuro di potassio, rende solubili l'oro e l'argento. La chimica delle reazioni tra il minerale e il cianuro è molto complessa, perché oltre all'oro e all'argento reagiscono altri elementi e composti. Tenendo conto soltanto della reazione con l'oro, che trasforma il metallo nativo in sale solubile, l'equazione chimica è:

$$4Au + 8NaCN + O_2 + 2H_2O \rightarrow 4NaAu(CN)_2 + 4NaOH$$

La soluzione viene trattata con polvere di zinco, che fa precipitare l'oro:

$$NaAu(CN)_2 + 2\,NaCN + Zn + H_2O \rightarrow Na_2Zn(CN)_4 + Au + H + NaOH.$$

Il minerale viene polverizzato e selezionato con macchine di vario tipo; anche i processi meccanici, come quelli chimici, vengono modificati continuamente, al fine di ottenere un alto recupero del metallo, riducendo la mano d'opera; il trattamento avviene senza soste.

L'oro è molto diffuso in natura: l'acqua del mare, per esempio, ne contiene 0,065 grammi per tonnellata (6 parti per miliardo); poiché il volume dell'acqua dei mari e degli oceani è molto grande, anche la quantità totale dell'oro è enorme, ma a nessuno può venire in mente di estrarre l'oro dall'acqua dell'oceano, per lo meno non con le tecniche attualmente in uso. Un minerale d'oro, anche usando le tecniche più recenti e tenendo conto degli alti prezzi attuali, deve contenere in media una decina di parti per milione per tutto il giacimento e di 50-70 parti per milione nelle zone più redditizie: questi sono i dati per la miniera di Carlin. Queste quantità segnano il limite inferiore di un minerale d'oro, cioè di una roccia dalla quale l'oro può essere estratto con profitto. Il contrasto medio tra il contenuto in metallo di un giacimento marginale e la roccia non mineralizzata è di 2500; per esempio, una normale roccia granitica ha un contenuto d'oro di 4 parti per miliardo, mentre il limite inferiore di un giacimento sfruttabile è un contenuto di metallo di 10 parti per milione; il contrasto (cioè il rapporto tra giacimento e roccia normale) è 10 parti per milione diviso 4 parti per miliardo, cioè 2500. In altre parole, la roccia di un giacimento sfruttabile contiene in media 2500 volte più oro delle rocce non mineralizzate. R. Kerrich e B. Fryer hanno riportato recentemente fattori di concentrazione di circa 15 000 in alcune miniere canadesi.

Si possono distinguere cinque tipi principali di giacimenti auriferi (i nomi dei tipi fanno riferimento in genere a una miniera esemplare, oppure alla prima in cui è stato identificato quel tipo): il tipo Mother Lode (filoni di quarzo aurifero), il tipo Homestake Mine (giacimenti di sostituzione selettiva), il tipo Witwatersrand (oro in lamelle disperso in silice), il tipo *placer* (giacimenti alluvionali) e il tipo Carlin (oro in granuli o lamelle submicroscopici, invisibili).

Presenteremo qui una breve descrizione di ciascun tipo di giacimento, per avere una visione d'insieme dei vari tipi di miniere d'oro. Per evitare ambiguità, è opportuno però chiarire il significato di alcuni termini. Prima di tutto, che cosa si intende con il sostantivo «minerale». I significati possibili sono due: il primo indica un elemento o un composto chimico di composizione definita, normalmente con una sua tipica forma cristallina. Per esempio, la pirite è un solfuro di ferro (FeS_2), di colore metallico brillante, giallo sul tipo dell'ottone; cristallizza in cubi, talvolta di notevoli dimensioni. La pirite è detta anche «l'oro degli sciocchi», perché gli inesperti la confondono spesso con l'oro, a causa del suo colore. Del resto, spesso la pirite è associata all'oro.

In un secondo senso, la stessa parola è usata per indicare la roccia, quasi sempre eterogenea, che viene sfruttata in un giacimento minerario. I francesi usano due parole distinte: *mineral* nel primo significato, *minerai* nel secondo. In italiano, si usa il termine «giacimento minerario», che indica l'insieme dei minerali dai quali è possibile estrarre con profitto una sostanza utile. Il termine è quindi legato non soltanto alla natura del minerale, ma anche alla quantità, al costo di estrazione e alle condizioni di mercato del minerale che si sfrutta.

Per quel che riguarda la forma in cui l'oro si presenta in un giacimento, adotterò qui questi termini: «oro visibile», quando l'oro si presenta disseminato nella roccia sotto forma di pepite, pagliuzze o granuli visibili a occhio nudo: «oro fine», quando l'oro può essere visto con una lente o sotto il microscopio, cioè si presenta in granuli o lamelle di dimensioni maggiori di due micrometri; «oro invisibile», quando i corpuscoli d'oro hanno dimensioni inferiori a due micrometri e possono essere studiati soltanto usando il microscopio elettronico e, naturalmente, con le indagini chimiche.

Nei primi anni della corsa all'oro in California, i piccoli gruppi di minatori usavano utensili rudimentali: la bateia, la pala, il piccone. Spesso scavavano pozzi e gallerie poco profondi: è qui rappresentato un pozzo azionato a mano. Anche questo disegno è tratto da *Geology of Northern California*.

La Sierra Nevada della California è costituita da una serie di plutoni granitici, che si sono intrusi in rocce metamorfiche più antiche nell'arco di un lungo periodo di tempo. Le più antiche intrusioni delle diverse masse di rocce fuse o semiliquide sono state datate a circa 200 milioni di anni fa, e quindi appartengono al Triassico; le più recenti sono datate a 80-90 milioni di anni fa, cioè al Cretaceo.

La Sierra Nevada è una grande catena montuosa tra la California e il Nevada, larga da 80 a 130 chilometri e lunga circa 650. Culmina al Mount Whitney (4650 metri) che fa parte di uno spettacolare panorama montuoso, con cime oltre i 4000 metri. Nella sua parte settentrionale, il batolite è fiancheggiato a ovest da una fascia metamorfica molto deformata, di età paleozoica e mesozoica. Il Mother Lode (che significa «filone madre») attraversa il cuore di questa fascia.

L'oro è contenuto in filoni di quarzo ed è generalmente il solo metallo che abbia tenori economici, ma è accompagnato da piccole quantità d'argento. I filoni di quarzo aurifero sono allineati in maniera discontinua su una lunghezza di quasi 200 chilometri, mentre le singole vene possono raggiungere una lunghezza di qualche chilometro e una larghezza di qualche metro; i filoni sono in massima parte verticali e sono stati lavorati in profondità fino a 1500 metri.

Durante le ultime fasi della complessa e lunga attività magmatica che ha originato il grande batolite, i fluidi idrotermali migrarono verso la superficie lungo le faglie e le altre fessure della roccia e raffreddandosi depositarono il quarzo e varie altre specie minerali, tra le quali anche l'oro e l'argento. Le vene e i filoni di quarzo sono molto numerosi, ma soltanto una piccola parte di essi contiene oro in quantità economica, con tenori superiori a 20 grammi per tonnellata.

Quasi tutte le miniere del Mother Lode sono abbandonate da decenni, ma il numero di quelle che furono attive nel passato è molto alto: molto spesso si trovano qua e là nelle montagne cunicoli, pozzi, o semplici scavi; le gallerie hanno perso la loro armatura di legno. Migliaia di chilometri di gallerie formano un intrico sotterraneo e raggiungono talvolta profondità superiori ai mille metri, come per esempio la Kennedy Mine (ora parco nazionale), che penetrava fino a 1500 metri e aveva una rete di gallerie lunga 200 chilometri.

La Homestake Mine, nello stato americano del South Dakota, è stata scoperta nel 1877 e venne acquistata e messa in produzione da George Hearst, il celebre magnate della prima ferrovia transcontinentale americana. Modesti depositi auriferi tipo placer erano già stati scoperti e sfruttati precedentemente a circa 90 chilometri dalla miniera di Homestake. L'oro veniva separato dal minerale per amalgama con il mercurio, processo con il quale si ricuperava soltanto una parte dell'oro; in seguito, si andò sviluppando il processo al cianuro, con il quale oggi si estrae normalmente il 97 per cento dell'oro. Insieme all'oro si produce anche l'argento, che ammonta in peso a circa un quinto dell'oro.

La formazione più produttiva è stata chiamata formazione di Homestake: uno scisto sideroplesitico-quarzitico (la sideroplesite è un carbonato di ferro e di magnesio di formula $FeCO_3\,MgCO_3$). Lo scisto è una roccia metamorfica che si divide facilmente in lastre sottili, come nelle ardesie, che sono una forma differente di scisti silicei. La formazione di Homestake, che è stata alterata dall'azione idrotermale, è del Precambriano ed è stata intrusa da anfiboliti (rocce a scarso contenuto di silice) alla fine del Precambriano (cioè oltre 570 milioni di anni fa) e da rocce ricche in silice di composizione granitica e sienitica del Terziario, circa mezzo miliardo di anni più tardi.

Abbiamo visto che le zone mineralizzate si trovano specialmente nella formazione di Homestake, che è uno scisto a sideroplesite e quarzo. Le specie minerali

Il blocco granitico della Sierra Nevada. Il batolite si allunga da sud a nord; ha pendenze relativamente dolci verso la Great Valley della California a ovest. A est precipita per faglia in una profonda fossa, il graben limitato dalla Sierra Valley e dalla Owens Valley. Il graben è costellato da bellissimi laghi, tra i quali il più esteso, e per molti il più bello, è il lago Tahoe. Verso est il graben è limitato da un'altra catena montuosa parallela, meno elevata, ma con molte cime al di sopra dei 3000 metri. Le cime più alte della Sierra Nevada superano i 4000 metri.

PAESE	TONNELLATE METRICHE
SUDAFRICA	703
URSS	325
CANADA	54
USA	36,4
TOTALE PARZIALE	1118,4
ALTRI PAESI	426,6
TOTALE	1545,0

La tabella riporta le stime, per l'anno 1980, della produzione di oro in tutto il mondo.

Geologia della Sierra Nevada, in California, da *Geology of Northern California*. Le diverse colorazioni indicano l'età delle rocce in milioni di anni prima del presente, precisata nella legenda.

dello scisto sono state sostituite da quarzo, clorite e arsenopirite per l'azione idrotermale connessa alle intrusioni di magma; nella aree più mineralizzate, lo scisto è stato completamente trasformato in una roccia densa di colore verde scuro o nero, costituita essenzialmente da clorite o da biotite (una mica) ricristallizzata; l'arsenopirite è disseminata nella roccia in modo non uniforme.

Questo primo stadio di sostituzione dei minerali e i due successivi sono caratterizzati dal fatto che i nuovi minerali non hanno riempito fessure o fratture preesistenti, nelle quali i fluidi idrotermali possono depositare le sostanze in soluzione. Le vene di alterazione sono dovute al rimpiazzamento delle specie minerali esistenti prima delle intrusioni magmatiche; ripetiamo che queste vene non sono riempimenti di spazi aperti, sebbene si trovino in zone di dilatazione delle rocce.

Il quarto e ultimo stadio di mineralizzazione è connesso alle intrusioni magmatiche di età terziaria che hanno mineralizzato specialmente la formazione Deadwood, del Paleozoico più antico.

Ognuno di questi processi geochimici può aver messo in posto l'oro nelle zone adatte.

Riassumendo, la miniera di Homestake ha due formazioni aurifere principali: la formazione Homestake e la formazione Deadwood. La prima è costituita da rocce sedimentarie metamorfiche del Precambriano, datate 2,5 miliardi di anni fa. Essa si è mineralizzata nel Cambriano, il periodo più antico del Paleozoico, per sostituzione selettiva.

La formazione Deadwood è del Cambriano ed è stata mineralizzata, anch'essa per sostituzione selettiva dei minerali, dall'attività idrotermale connessa a intrusioni magmatiche del Terziario: le due mineralizzazioni sono separate nel tempo da circa 500 milioni di anni!

Nella miniera di Homestake, l'oro è visibile a occhio nudo soltanto in casi eccezionali; come norma, le particelle d'oro si vedono bene al microscopio; i geologi di lingua inglese parlano in questo caso di *fine gold,* che si può tradurre in italiano con «oro microscopico». Il lettore si domanderà come mai la miniera sia stata scoperta da quei rozzi cercatori, che certamente non andavano sul terreno con il microscopio. La risposta è semplice: il minerale veniva trattato con mercurio per ottenere l'amalgama d'oro e mercurio. L'amalgama veniva riscaldato in apposite storte, dove il mercurio volatilizzava, per essere poi ricuperato sotto forma liquida per raffreddamento. Nell'antichità, l'amalgama veniva filtrato in una sacca di pelle di camoscio, separando così i due metalli.

Le formazioni aurifere della miniera di Homestake contengono circa 100 parti per milione d'oro.

Nel tipo Witwatersrand, il minerale è costituito da un conglomerato di ciottoli di quarzo, ognuno dei quali ha un diametro di qualche centimetro. Questi ciottoli sono saldati insieme da un cemento siliceo. L'oro è mescolato a pirite ed è disperso nel cemento siliceo. Le lamelle d'oro sono molto piccole: hanno appena qualche micrometro di lunghezza e quindi sono invisibili a occhio nudo. Per fare un esempio, hanno dimensioni dell'ordine di quelle dei batteri e quindi, come i batteri, sono visibili al microscopio normale, il cui limite di visibilità è di due micrometri.

Il giacimento principale è quello di Witwatersrand. I tenori medi sono bassi (da 7 a 12 parti per milione), ma aumentano considerevolmente in un conglomerato a ciottoli più grossi, il Main Reef. In realtà non si tratta di *reef* nel significato proprio del termine, che si usa per indicare la scogliera corallina: il vocabolo viene usato nel Witwatersrand per i conglomerati auriferi, perché questi, estremamente duri, sono stati erosi dagli agenti atmosferici molto meno degli strati che li accompagnano, e quindi sporgono al di sopra del terreno circostante. I conglomerati grossolani comprendono molti strati, ognuno dello spessore di parecchi metri *(pay streak).* Il reef di Witwatersrand si allunga per 180 chilometri; altri reef si trovano nel Transvaal e nell'Orange.

L'origine del giacimento è oscura. In un primo tempo, si è pensato che si trattasse di un placer fossile, e si avevano buoni argomenti a favore di questa ipotesi. Infatti, il conglomerato aurifero è chiaramente un deposito alluvionale; inoltre, la serie stratigrafica si mantiene costante anche su lunghe distanze, indipendentemente dai filoni di diabase o da faglie che l'attraversano. D'altra parte, l'oro si trova entro i cristalli di pirite, o sulla loro superficie; per di più, è associato a molti altri minerali, come pirrotite, cobaltite e miche, generalmente di altra origine. Inoltre, l'oro penetra talvolta anche nelle minutissime fessure dei ciottoli del conglomerato. Si pensa ora che Witwatersrand sia un antico giacimento d'oro alluvionale, che in un'epoca geologica più recente subì l'azione di soluzioni idrotermali. In altre parole, i conglomerati sarebbero stati mineralizzati una prima volta nel Paleozoico e una seconda volta nel Terziario, durante i periodi di raffreddamento di due diverse intrusioni granitiche. Secondo l'espressione di Raguin, «il giacimento è un placer metamorfizzato nel corso dei tempi».

I giacimenti del Rand sono sfruttati con macchinari e tecniche molto moderni. Varie centinaia di migliaia di lavoratori negri lavorano nel sottosuolo (fino a 2500 metri di profondità) e in superficie. Fino a che il prezzo dell'oro è stato fermo a 35 dollari all'oncia, le miniere erano redditizie soltanto perché le paghe dei minatori erano basse; oggi, con il prezzo oscillante intorno ai 300 dollari l'oncia troy (circa 10 dollari al grammo), l'Unione Sudafricana, che è il maggior produttore d'oro del mondo, ha entrate molto cospicue.

I giacimenti tipo placer sono molto comuni; in linea generale, la quantità d'oro di una singola area non è grande, anche se la produzione complessiva può raggiungere cifre degne di nota. Tanto

Zone aurifere della Sierra Nevada. Nell'area occidentale della fascia aurifera venne scoperto per caso il primo giacimento da alcuni operai che scavavano un canale di alimentazione per un mulino ad acqua. Cominciò così, nel 1848, la corsa all'oro della California. Per circa trent'anni si ottennero produzioni ingenti, poi, a poco a poco, i giacimenti redditizi si esaurirono e la ricerca dell'oro si spostò quasi completamente in altri stati. Ora l'esplorazione ha assunto una forma diversa: nel Nevada e in California si sono scoperte diverse miniere di oro invisibile. Il singolo gruppetto di pochi uomini ha ceduto il passo alle grandi compagnie minerarie, che impiegano ingenti capitali nell'esplorazione e nello sfruttamento meccanico delle miniere.

Struttura delle rocce precambriane nella zona della miniera di Homestake. La formazione Homestake è indicata in colore. Le altre formazioni sono indicate dalle sigle: Poorman, PMF; Ellison, EF; Northwestern, NWF; Flagrock, FRF; Grizzly, GF. I grandi ammassi intrusivi del Terziario sono in grigio scuro. I grandi ammassi anfibolitici sono in grigio chiaro. Le faglie sono indicate da linee tratteggiate. Non sono indicati invece gli strati del Cambriano, i piccoli corpi intrusivi del Terziario, i piccoli corpi anfibolitici e i conglomerati del Terziario. La colonna stratigrafica è formata dall'alto, come segue: conglomerati terziari, quaternari e recenti (almeno 30 metri); discordanza; rocce intrusive terziarie; formazione Deadwood del Cambriano (100-160 metri); discordanza; rocce intrusive precambriane; rocce sedimentarie precambriane: formazioni Grizzly (1000 metri o più), Flagrock (1500 metri); discordanza; Northwestern (1300 metri), Ellison (1000-1500 metri), Homestake (70-100 metri), Poorman (700 metri o più). L'illustrazione è rielaborata a partire da un originale di A. L. Slaughter della Homestake Mining Company.

per citare un esempio, i placer dell'Oregon hanno contribuito alla produzione totale dello stato per più della metà.

L'origine di questi depositi è semplice: le acque dei torrenti e dei fiumi che attraversano un'area tipo Mother Lode (o qualsiasi altra area dove il minerale aurifero affiora in superficie) erodono il minerale e lo trasportano a valle. In molti punti l'acqua rallenta la sua corsa verso il piano e deposita i sedimenti che ha trasportato a valle: ciottoli, sabbie e argilla. L'oro, che è più pesante (peso specifico di 19,33, mentre il peso specifico del quarzo e del granito è di circa 2,3 -3,5), deposita per primo, e si forma così un giacimento secondario. In linea generale, quindi, i placer sono di età recente, perché altrimenti sarebbero stati erosi. In certi casi, però, il placer originario è stato protetto dall'erosione da altre rocce, che si sono depositate al di sopra dei ciottoli o delle sabbie del giacimento fluviale. Tanto per dare un esempio di questo tipo di giacimenti «salvati dalle acque», citiamo ancora le pendici occidentali della Sierra Nevada: più in alto nella Sierra, vi è la lunga fascia del Mother Lode, che affiorava già nel periodo Terziario e veniva erosa dai numerosi torrenti che scorrevano a ovest, verso il mare che allora occupava la Great Valley della California. Durante l'Eocene (circa 50 milioni di anni fa), grandi eruzioni vulcaniche copersero di lava una vasta regione; gli antichi fiumi cambiarono il loro corso e i sedimenti fluviali non subirono più l'erosione, che riprese soltanto in un tempo prossimo al nostro. Dove i torrenti sono riusciti a incidere una valle, erodendo la lava, i vecchi ciottoli auriferi rividero la luce, dopo tanti milioni di anni. Questi giacimenti vennero lavorati con nuove tecniche. I minatori disgregavano il conglomerato di ciottoli e le altre rocce con potenti getti d'acqua, che veniva incanalata in manufatti, talvolta molto grandi, atti a separare l'oro. Più comuni di questi placer fossili sono quelli recenti, formatisi durante l'attuale ciclo erosivo.

Molti dei fiumi che erodono i massicci granitici trasportano piccole quantità d'oro. In Italia, sono aurifere le sabbie di molti fiumi piemontesi, specialmente quelli che scendono dalla catena del Monte Rosa. Sono invece filoniani (del tipo Mother Lode) i giacimenti delle miniere di Pestarena, presso Macugnaga, e di Feuillaz in Val d'Aosta. Queste miniere, di modestissima importanza, sono state abbandonate molti anni fa. Ricordiamoci però che i fenici e gli antichi greci consideravano l'Italia una regione aurifera al tempo di Omero (700-800 anni a.C.). Plinio, poi, riporta che nel primo secolo d.C. era proibito impiegare nel Vercellese più di 5000 uomini nel lavaggio delle sabbie aurifere.

La miniera di Carlin, nello stato del Nevada, fu la prima miniera sfruttante un giacimento aurifero contenente soltanto oro invisibile. Nel distretto aurifero di Lynn, al quale appartiene Carlin, l'oro fu scoperto per la prima volta nel 1907. In seguito, per molti anni, si ebbe una produzione sporadica da numerosi piccoli giacimenti tipo placer e da scisti silicei del Paleozoico (formazione Vinini dell'Ordoviciano). Il minerale aurifero veniva scavato in superficie o a profondità modesta, con scarso successo. L'oro era visibile e quindi facilmente recuperabile anche con la bateia. I lavori successivi dimostrarono che anche a piccola profondità non esiste oro invisibile. Nel 1960 venne pubblicato un rapporto di un geologo dell'US Geological Survey, R. J. Roberts, che descriveva gli allineamenti di rocce osservabili nel distretto di Lynn. La compagnia mineraria Newmont prese nota del rapporto e già l'anno seguente mandò i suoi geologi a rilevare la zona. Questo lavoro di terreno portò all'individuazione di una zona ritenuta favorevole. I diritti minerari appartenevano a un vecchio minatore, che per anni aveva cercato l'oro con i metodi tradizionali, ma non era mai riuscito a vedere «colori», cioè, nel gergo minerario, non aveva mai potuto vedere al fondo della sua bateia la minima pagliuzza del metallo in quell'area. Egli fu ben contento di vendere i suoi diritti per duemila dollari a «un tizio di passaggio», che era, naturalmente, un agente della Newmont.

I lavori di esplorazione procedettero velocemente: la Newmont perforò un grande numero di pozzi carotati: già il terzo pozzo mostrava all'analisi l'oro invisibile con tenori industriali a piccola profondità. In totale furono perforati, carotati e analizzati complessivamente circa 200 000 metri di roccia. I pozzi dimostrarono che lo strato mineralizzato aveva uno spessore variabile da qualche metro a trenta metri. La profondità del giacimento si aggirava sui 130 metri con un volume complessivo di 11 milioni di tonnellate e con un tenore medio di 0,32 once troy per tonnellata, il che equivale nel sistema metrico a 10 parti per milione. Poiché nell'industria mineraria e nel commercio dell'oro si continua a usare il sistema inglese di misure, diamo i valori equivalenti nel sistema metrico. Una oncia troy (da non confondere con l'oncia avoirdupois) corrisponde a 31,1 grammi. Una tonnellata avoirdupois equivale a 1016 chilogrammi. Quindi 0,32 once troy per tonnellata sono 9,95 grammi d'oro per 1 016 000 grammi, cioè, arrotondando, 10 grammi d'oro ogni milione di grammi (una tonnellata) di minerale, quindi circa 10 parti per milione. In conseguenza, la prima valutazione della quantità totale d'oro indicava la cifra di 3,52 milioni di once troy, cioè, arrotondando, 110 tonnellate metriche di metallo. Al prezzo d'allora (35 dollari all'oncia, cioè 1,10 dollari al grammo), il valore complessivo dell'oro era dunque stimato a 120 milioni di dollari.

La Newmont decise di esplorare la zona di Carlin nel 1961; appena quattro anni dopo, nel maggio 1965, la prima carica d'oro della fornace di fusione versava un ruscello rutilante d'oro puro.

Per giungere a questo risultato, la compagnia aveva speso dieci milioni di dollari; ma già alla fine del primo anno di operazioni regolari Robert Fulton, che aveva diretto la difficile e complessa realizzazione, comunicò che la produzione dell'anno ammontava a sette milioni di dollari.

A cinque anni dall'inizio, Carlin lavorava 2500 tonnellate di minerale al giorno e ricavava annualmente una quantità d'oro che, al prezzo attuale oscillante attorno ai 300 dollari l'oncia, corrisponde a circa 90 milioni di dollari.

La scoperta dell'oro invisibile a Carlin indusse altre compagnie a intraprendere l'esplorazione sistematica di una vasta zona lungo lo stesso allineamento geologico. Furono così scoperte le miniere di Gold Acres, Cortez, Bootstrap e Getchell, tutte nelle Roberts Mountains, come Carlin. L'esplorazione continua nel Nevada e in altri stati americani. Le compagnie tengono naturalmente segreti i risultati, ma talvolta filtra qualche notizia non ufficiale. Si dice che una compagnia abbia individuato nel Nevada un giacimento dieci volte più importante del giacimento di Carlin. Inoltre, la Homestake ha comunicato ufficialmente la scoperta di un giacimento in California, a una distanza di circa cento chilometri a nord di San Francisco.

Abbiamo già parlato dell'abbondanza dell'oro in natura, ma a questo punto della discussione è utile fornire qualche cifra supplementare. In media, le rocce della crosta terrestre contengono 0,004 parti per milione di oro; le differenze tra questa cifra media e quella dei vari tipi di rocce (naturalmente senza considerare i giacimenti sfruttabili) sorprendono per la loro piccola entità: 0,005 parti per milione per le rocce ultramafiche e i calcari; 0,004 per i basalti, le granodioriti, i graniti e le argille; le acque dei fiumi contengono soltanto 0,002 parti per milione di oro, cioè la metà della media della crosta terrestre.

Il problema geochimico dell'origine dei giacimenti auriferi consiste quindi nell'individuare i processi naturali atti a trasportare e a concentrare l'oro, da rocce che ne contengono quantità insufficienti, ad altri ambienti geologici.

Nei giacimenti secondari tipo placer il problema è semplice, come si è visto, perché si parte da rocce già ricche in oro; i semplici processi dell'erosione e della concentrazione per gravità, dovuti all'azione delle acque superficiali, spiegano l'origine di questo tipo di giacimenti.

Per i giacimenti primari, le cose sono molto più complicate; la comprensione di alcuni processi geologici e geochimici legati alla formazione dei giacimenti auriferi è dovuta a ricerche iniziate e perseguite durante gli ultimi 25 anni; questo intenso lavoro scientifico ha portato a nuove teo-

La miniera di Homestake fa parte dell'importante distretto minerario dei Black Hills; si ritiene che nel distretto, di cui è data qui la struttura geologica, si possano trovare nuovi giacimenti.

rie sulle intrusioni magmatiche, sul trasporto e la concentrazione dei minerali e soprattutto sui movimenti e le caratteristiche geochimiche delle soluzioni acquose calde. Malgrado tanti progressi, molti dubbi e molti punti oscuri offuscano ancora la visione d'insieme; come conseguenza, il progresso nei metodi di esplorazione non è ancora neanche lontanamente paragonabile a quello raggiunto nella ricerca del petrolio, che è alla vetta della sofisticazione scientifica e tecnica; anche altri settori minerari sono ben più avanzati, come dimostrano molte scoperte spettacolari di nuovi giganteschi giacimenti di altri metalli. Bisogna però aggiungere subito che la scoperta e lo sviluppo industriale dei giacimenti d'oro invisibile ha trasformato quest'industria in modo profondo.

L'esplorazione aurifera non è più limitata al colpo di fortuna del ricercatore primitivo, talvolta un po' aiutato dall'intuizione dei geologi. La scoperta di Carlin ha iniziato una nuova fase: l'esplorazione con metodi scientifici, al di là del potere dei sensi.

I fatti principali che conosciamo sono già stati esposti in parte nelle pagine precedenti; cercherò ora di riassumerli in modo più sistematico, anche se lo spazio mi obbliga ad accennare soltanto in modo molto sommario ad alcune importanti ricerche teoriche di geochimica, in particolare a quella di Helgeson e dei suoi allievi, che lavorano all'Università della California a Berkeley.

Cominciamo dalle più ovvie osservazioni geologiche, già intuite o conosciute da molto tempo.

Nella miniera di Homestake, la roccia veniva perforata a mano fino al 1900; attualmente si usano 120 milioni di chilowattora per anno per azionare le innumerevoli macchine della miniera. Nella fotografia, è visibile una trivella ad aria compressa. Nel foro ottenuto, si fa poi brillare l'esplosivo.

Le pareti delle gallerie nella miniera di Homestake sono rinforzate, ma l'armatura di legno ha dei passaggi per l'acqua che penetra nel terreno. Essa viene quindi sollevata fino alla superficie.

I giacimenti auriferi possono trovarsi in ogni tipo di roccia, come già appare dalla descrizione precedente dei tipi di giacimenti. Si conoscono giacimenti incassati in rocce plutoniche, altri entro rocce vulcaniche, o sedimentarie, o metamorfiche, appartenenti a età diversissime. L'età e la petrologia delle rocce incassanti sono di interesse locale, ma non ci possono dire molto sull'origine dei giacimenti.

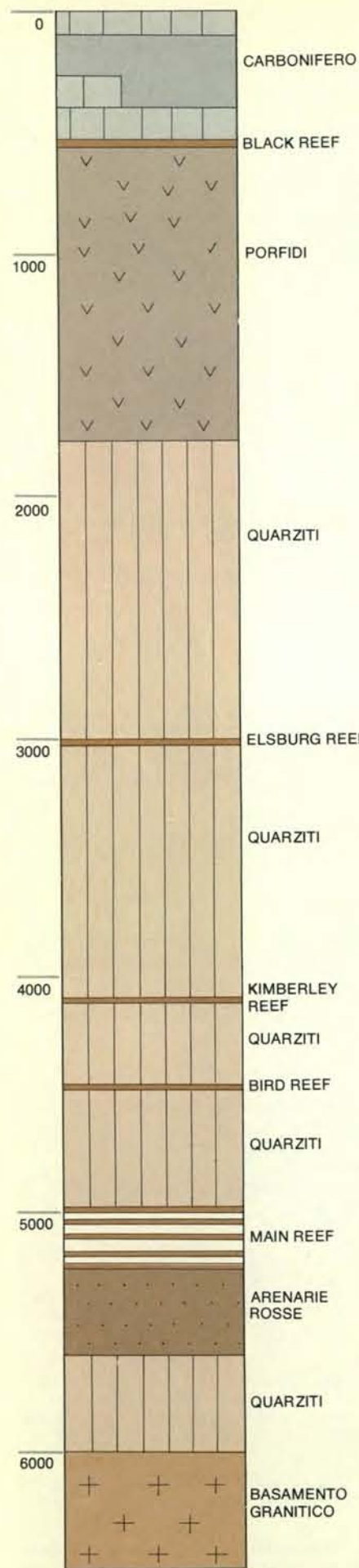
Ciò che sappiamo di sicuro è che l'oro proviene in ultima analisi da rocce magmatiche. Infatti, essendo un elemento pesante, ben difficilmente poteva far parte della crosta terrestre durante i primi stadi della sua formazione. L'oro doveva essere più abbondante nel mantello, cioè la parte della Terra che sta tra il suo denso nucleo di ferro e la crosta, dalla quale è separato dalla discontinuità sismica di Mohorovičić. Il mantello è composto da materiali simili a quelli delle meteoriti ferrose e quindi la sua composizione è eterogenea. Fino a circa trent'anni fa, le conoscenze sul mantello erano molto scarse: si credeva allora che la sua composizione fosse uniforme e omogenea e che la sua partecipazione ai grandi processi geologici della crosta fosse trascurabile: il mantello era considerato inerte e poco interessante. Ma gli studi recenti hanno fatto piazza pulita di quelle vedute. I progressi relativi agli strumenti e ai metodi geofisici, lo studio geologico sistematico degli oceani e la nuova visione dell'assetto generale dei continenti, progressi che tutti insieme hanno portato alla rivoluzione scientifica della tettonica a zolle, hanno rivelato che il mantello è eterogeneo e che si muove, lentamente, ma continuamente. Il mantello forma nuova crosta terrestre nelle grandi rughe da cui sgorga la lava nel fondo degli oceani, mentre riassorbe al margine dei continenti pezzi della «vecchia» crosta, trasportati al di sotto della discontinuità di Mohorovičić. I grandi fatti geologici a scala planetaria, come il vagabondare delle zolle continentali, i movimenti delle rocce fuse nella crosta terrestre e l'incessante formazione e distruzione delle montagne sono dovuti in ultima analisi all'attività del mantello.

Il fenomeno geologico che spiega la migrazione degli elementi pesanti dal mantello alla crosta terrestre è indicato dal termine generale «attività ignea», che designa tutti i processi inerenti alla fusione delle rocce, cioè la trasformazione da roccia a magma. L'attività ignea determina due specie di processi: il vulcanismo, che consiste nell'eruzione in superficie di magma, e il plutonismo, cioè l'intrusione di magma profondo nelle rocce sovrastanti, dove si raffredda e si solidifica. La composizione chimica delle rocce derivate dal raffreddarsi delle lave vulcaniche e quella delle rocce plutoniche raffreddatesi in profondità è molto simile, mentre molto diverso è il loro aspetto; il magma che le origina è costituito da silicati fusi, nei quali la silice ha in generale valori compresi tra il 45 e il 70 per cento in peso. I magmi e le rocce che ne derivano si dicono «acidi» quando hanno un alto contenuto di silice e «basici» quando il contenuto in silice è basso. Una classificazione molto semplificata delle rocce ignee comprende tre tipi principali di magma e di rocce vulcaniche e plutoniche che ne derivano: «magma acidi» (per esempio le rocce granitoidi e i graniti nella serie intrusiva, rocce riolitiche in quelle effusive); «magma intermedi», con contenuto in silice attorno al 60 per cento, dai quali derivano per esempio le andesiti e le daciti nella serie effusiva, le dioriti e le sieniti in quella intrusiva; «magma basici» (silice 45-60 per cento circa) da cui derivano le rocce effusive basaltiche e le rocce intrusive ultrafemiche, le peridotiti e i gabbri.

Le rocce plutoniche provenienti dal mantello hanno trasportato l'oro nella crosta terrestre e devono quindi essere considerate l'origine prima e lontana dei giacimenti auriferi. L'oro è diffuso nel magma; esso deve uscirne per essere concentrato e questo avviene per azione delle soluzioni acquose calde. Poiché la quantità media dell'oro è circa la stessa tanto nelle rocce magmatiche quanto in quelle sedimentarie, i giacimenti possono essere dovuti alla concentrazione dell'oro contenuto in qualsiasi tipo di roccia, anche in tenori inferiori alla media (4 parti per miliardo). Nei giacimenti idrotermali l'oro non è mobilizzato dall'acqua, ma da soluzioni acquose calde. E poiché il contenuto in acqua del magma è molto basso, attorno al 5 per cento, le rocce fuse tendono ad assorbire l'acqua piuttosto che a lasciarla sfuggire verso le rocce incassanti. Ciò è stato definitivamente provato negli ultimi 25 anni mediante lo studio degli isotopi dell'idrogeno e dell'ossigeno e per via sperimentale ed è ormai un punto fermo della geochimica. Ma queste soluzioni devono essere calde e per questo si parla di fenomeni idrotermali, usando una parola composta dal greco.

La temperatura dell'acqua circolante nelle rocce è dunque il fattore più importante nell'origine dei giacimenti auriferi e argentiferi, come di altri, specialmente di solfuri (rame, piombo, zinco, uranio, ferro ecc.). Perché calde? La risposta a questa domanda, formulata in termini così generali, è molto complicata; ma, per quanto ci interessa, è sufficiente esporre alcuni fatti generali ormai ben noti, almeno nelle linee generali e in modo semiquantitativo.

Il lavoro della miniera richiede un continuo rifornimento di aria fresca anche nei livelli più profondi. Inoltre, occorre pompare fino alla superficie le grandi quantità di acqua che percolano attraverso la roccia. Si vedono in fotografia installazioni di pompaggio della miniera di Homestake.

Una gabbia degli ascensori che trasportano uomini e minerali a Homestake. Le quattro fotografie di queste pagine sono riprodotte per gentile concessione della Homestake Mining Company.

In generale, la temperatura della crosta terrestre cresce con la profondità. Fino a poche decine di anni fa, si disponeva di ben poche misure utilizzabili della temperatura in profondità, dove le variazioni annue della temperatura del suolo non disturbano più, almeno nelle aree dove le acque superficiali si muovono molto lentamente. Oggi le cose sono cambiate, perché la quantità di dati va aumentando rapidamente. Possiamo ormai utilizzare decine di migliaia di misure tanto del gradiente di temperatura quanto del flusso di calore. Limitiamoci a parlare in modo sommario del gradiente: la media di tutte le misure indica un gradiente di 3 gradi centigradi per cento metri. Vi sono però aree fredde, dove il gradiente è minore e può essere perfino negativo (e altre dove il gradiente supera 1 grado centigrado per metro), almeno per alcune centinaia di metri di profondità. Il calore dei fluidi prodotti da pozzi perforati nelle aree calde può essere ed è sfruttato per produrre elettricità e per molti altri usi; questa speciale attività mineraria è l'oggetto dell'industria geotermica, che ha rivelato molti fatti importanti per la comprensione della genesi dei giacimenti idrotermali.

Due sono le sorgenti principali che forniscono calore alla crosta terrestre. La più importante è la conduzione di calore dal mantello di rocce sottostanti alla discontinuità di Mohorovičić, dove la temperatura è di circa 1000 gradi centigradi a 200 chilometri di profondità nelle zone centrali dei continenti e a 900 chilometri nelle corrispondenti zone degli oceani. Il calore del mantello è fornito dall'attività radioattiva del mantello superiore solido e degli altri involucri dell'interno della Terra, ivi compreso il nucleo, per il quale si valuta una temperatura di 6000 gradi centigradi al centro della Terra.

Poiché le rocce del mantello superiore hanno una temperatura vicina al punto di fusione, una diminuzione della pressione permette il passaggio dallo stato solido a quello liquido, con la conseguente mobilizzazione del magma verso gli strati superiori della crosta terrestre. A loro volta, i fluidi contenuti in quest'ultima si riscaldano ulteriormente, aumentando così la loro capacità di dissolvere le sostanze delle rocce attraverso le quali circolano, e ne risultano le soluzioni idrotermali. Queste, circolando nelle rocce incassanti, finiscono con il raffreddarsi; una parte delle sostanze disciolte si cristallizza e si deposita, generando talvolta i giacimenti minerari, quando questi processi geochimici avvengono nell'ambiente geologico adatto: il loro volume è piccolo, poiché anche i più grandi occupano meno di 4 chilometri cubi. Come abbiamo già visto, anche i giacimenti auriferi primari appartengono alla grande famiglia dei giacimenti idrotermali.

Serie stratigrafica del giacimento di Witwatersrand, Sudafrica. I reef sporgono dal terreno circostante perché sono molto duri e quindi meno soggetti all'erosione. Questi conglomerati costituiscono gli strati più mineralizzati.

L'oro è poco solubile e forma molto difficilmente composti. Spesso si trova allo stato nativo o sotto forma di lega con altri metalli, specialmente l'argento. È spesso associato a giacimenti di altri minerali, specialmente quelli di solfuri di ferro e di rame, sotto forma di oro nativo. Un'altra parte notevole di tutto l'oro estratto è un sottoprodotto dei giacimenti di solfuri.

Malgrado quest'associazione dell'oro ai giacimenti di solfuri, l'oro non è stato mai trovato in natura come solfuro, anche se i suoi cristalli sono cresciuti insieme con quelli di piriti e di altri solfuri, che hanno sostituito i cristalli di quarzo nelle sue vene. In questo caso, l'oro e i solfuri si sono depositati dopo il quarzo; ma in altre vene le tre specie minerali (oro, solfuri e quarzo) si sono cristallizzati nello stesso tempo. Questi e altri fatti dimostrano che l'oro è giunto al suo luogo di deposizione in soluzioni acquose calde ed è precipitato in particolare a seguito del loro raffreddamento. In certi casi, però, e in modo particolare quando la roccia incassante è stata molto alterata, un aumento dell'acidità delle soluzioni (cioè una diminuzione del *p*H) può essere un fattore altrettanto o più importante.

Helgeson ha calcolato l'ordine di grandezza della massa trasportata nella formazione di un giacimento d'oro di dimensioni rispettabili. Come esempio, per precipitare una vena di quarzo alta 300 metri, lunga 300 e spessa 30 centimetri, considerando che la soluzione idrotermale da 300 gradi centigradi si raffreddi a 200 gradi centigradi, occorrono 2,3 × 10^8 tonnellate di soluzione. Assumendo che la soluzione sia satura in oro, si depositano circa 4,3 tonnellate d'oro e 13 000 tonnellate di pirite insieme al quarzo. Assumendo inoltre un flusso di 280 metri cubi al giorno, basterebbero 2000 anni per formare il giacimento e la velocità di flusso della soluzione nella vena sarebbe di 0,2 centimetri al minuto.

Nelle soluzioni idrotermali, con temperature da 200 a 300 gradi centigradi, l'oro esiste sotto forma di complessi - principalmente di oro e tellurio, come calaverite ($AuTe_4$), silvanite ($Ag_2Au_2Te_8$), krennerite (Au_8Te_{16}) ecc. - e di colloidi, cioè sospensioni di particelle molto piccole, al di sotto di 2 micrometri, che possono essere rivestite da una pellicola di materiale organico.

Lo studio dei fenomeni geotermici ha fatto progredire notevolmente le nostre conoscenze su tutti i giacimenti idrotermali, compresi quelli d'oro. Alcuni depositi silicei o travertinosi formatisi attorno alle sorgenti calde contengono oro in concentrazioni notevoli; a Steamboat Springs (Nevada) D. White ha trovato 10 parti per milione di oro e 400 parti per milione di argento e A. Grange riporta 2 parti per milione di oro e 120 di argento in un deposito della Nuova Zelanda. Weissberg, effettuando un riassunto di tutti i dati relativi alle zone termali della Nuova Zelanda, è giunto a conclusioni che si accordano con i raffinati studi teorici di Helgeson.

Cerchiamo ora di riassumere i fatti principali sulla genesi dei giacimenti d'oro invisibile. Il susseguirsi dei processi è schematizzato dal seguente modello geologico-geochimico, che appare chiaro nelle grandi linee, mentre comprende molti punti poco noti nei particolari.

Una massa magmatica, presumibilmente granitica, si intrude nella crosta terrestre a una profondità di 8-15 chilometri. Le rocce nelle quali si intrude vengono alterate e diventano più o meno impermeabili, sia per la deposizione di minerali (specialmente silice) nelle fessure e fratture, sia per la trasformazione delle specie chimiche esistenti in altre. Contrariamente a ciò che si pensava fino a poche decine di anni fa, il magma non fornisce soluzioni acquose alla roccia incassante, ma avviene piuttosto il contrario: la roccia fusa assorbe fluidi dall'esterno.

Il magma però ha temperature di 500-800 gradi centigradi e quindi riscalda la roccia circostante e i fluidi in essa contenuti. Se la roccia è sufficientemente permeabile, i fluidi caldi, che sono meno densi, si muovono verso l'alto o lateralmente, seguendo le vie di maggiore permeabilità. Si forma così un sistema convettivo, con correnti che si muovono verso l'alto, compensate dalla discesa di fluidi più freddi. Il processo convettivo continua fino a che il magma fornisce calore in quantità sufficiente. I fluidi continuano a sciogliere le sostanze minerali, fino a divenire sature entro i limiti della temperatura, della pressione e delle quantità e qualità delle specie chimiche presenti nelle soluzioni. Infine, dove le correnti convettive incontrano una barriera impermeabile, la pressione e la temperatura sono inferiori e le soluzioni idrotermali depositano una parte delle sostanze solubili, riempiendo a mano a mano le fessure e gli altri spazi dove circolano.

Questo modello spiega a grandi linee i fatti essenziali relativi all'origine dei giacimenti idrotermali: il passaggio in soluzione delle specie chimiche che nelle rocce incassanti erano allo stato cristallino o amorfo; il loro trasporto verso l'alto; infine, la cristallizzazione e la concentrazione delle specie chimiche contro una barriera di impermeabilità o in superficie.

Oltre a questi processi di natura fisico-chimica, il modello spiega anche i più importanti fatti geologici legati ai giacimenti idrotermali: l'evidente vicinanza di masse magmatiche intruse; la loro erratica geometria; l'insediarsi della maggioranza di essi in profondità, che l'erosione può progressivamente diminuire, anche fino a riportarli alla luce.

La miniera di Carlin e le altre miniere vicine. Le «finestre» di Paleozoico carbonatico sono indicate dalle zone in colore. La figura è stata rielaborata da un lavoro di D. M. Hansen e Kerr.

J. G. Wargo, vicepresidente della San Francisco Mining Associates, ha fornito nuovi particolari interessanti sulla storia della scoperta della miniera di Carlin in un articolo pubblicato nel 1979 sulla rivista «Mining Engineering». Nello stesso articolo, Wargo anticipa l'evoluzione della ricerca di nuovi giacimenti dello stesso tipo. Per quanto si riferisce alla scoperta di Carlin, mi sembra opportuno citare direttamente dal suo testo: «Nel 1961 - scrive Wargo - Robert B. Fulton e John Livermore, della Newmont, seguendo un'idea ispirata da una pubblicazione dell'US Geological Survey, scelsero un'area dove affioravano strati geologici più antichi dei circostanti (finestra geologica), per nuovi lavori di ricerca. La scelta sembra essere stata fatta pensando che la finestra fosse dovuta a un sollevamento dovuto all'intrusione di rocce plutoniche in profondità». Le rocce intruse potevano aver originato la mineralizzazione.

«Dopo un periodo di rilevamento e di raccolta di campioni, si passò a un programma di perforazione e il terzo pozzo penetrò uno strato dello spessore di 20 metri, che all'analisi dava 34 grammi per tonnellata (1,0 once per short ton). Questo risultato segnò l'inizio della miniera di Carlin e di una corsa all'oro con lo scopo di ritrovare depositi dello stesso tipo.

«La scoperta di Carlin, e, pochi anni dopo, di Cortez, indussero a mettere l'accento sulla relazione spaziale, se non genetica, dell'oro con la faglia di sovrascorrimento delle Roberts Mountains, faglia legata a una delle più importanti strutture del Nevada centrale. L'erosione selettiva ha rimosso una parte delle rocce della falda superiore, mettendo in luce gli strati carbonatici della falda inferiore. Questa formazione era la roccia entro la quale si trovavano i minerali d'oro delle miniere di Carlin e di Cortez. Il primo modello grezzo, non raffinato, era perciò basato su: (1) la presenza della faglia di sovrascorrimento delle Roberts Mountains, e (2) la presenza di una finestra che espone in superficie rocce carbonatiche della falda inferiore.»

Questo rozzo modello guidò un certo numero di compagnie minerarie nella scelta di nuove aree di ricerca e quasi tutte le finestre furono esplorate perforando pozzi di ricerca. Alcune finestre risulta-

Stratigrafia della miniera di Carlin, nel Nevada. La successione stratigrafica, dall'alto al basso, comprende la formazione Popovich del Devoniano, con calcari stratificati, dolomitici, fossiliferi, con qualche strato di dolomite; una discordanza a silt e argille silicizzate, con barite; la formazione delle Roberts Mountains, con silt argillosi o dolomitici, talvolta mineralizzati, alterati e silicizzati. Lo strato mineralizzato (in grigio scuro) fa parte di questa formazione.

rono essere blocchi slittati per gravità, «una sorpresa per i geologi, cui pareva che i loro pozzi passassero dalle rocce della falda inferiore a quelle della falda superiore».

La scoperta di Carlin diede l'avvio a un programma federale di incentivi per la ricerca di metalli scarsi negli Stati Uniti, come l'oro, l'argento, il platino, lo stagno, l'antimonio, il bismuto, il nichel e il tantalio. In un primo tempo, si diede priorità alla ricerca dell'oro. Il Geological Survey e l'Ufficio Minerario pubblicarono numerosi studi, che misero in luce gli aspetti geologici e geochimici più importanti dei giacimenti tipo Carlin. Le analisi di campioni di rocce raccolti nel corso di studi precedenti indicarono un'importante anomalia d'oro e questi dati portarono subito alla scoperta del giacimento di Cortez, ora esaurito, che produsse 26 400 chilogrammi d'oro.

Un nuovo programma per i metalli «strategici», che naturalmente comprende anche l'oro, è ora in corso di approvazione da parte del Congresso degli Stati Uniti e costituisce un punto importante del programma governativo.

Le nuove informazioni sull'oro invisibile hanno permesso di formulare un modello genetico più raffinato, che è stato già esposto con alcune modificazioni nelle pagine precedenti. Wargo lo espone sinteticamente in questo modo: la mineralizzazione è dovuta a processi idrotermali guidati da strutture geologiche anteriori e la precipitazione dell'oro avvenne per ebollizione della soluzione idrotermale e per le reazioni della roccia ospite in presenza di minerali carbonatici e sulfurei. L'oro in granuli estremamente minuti si trova nei cristalli di pirite e di altri solfuri e con il carbone.

Il lettore osserverà che il modello da me adottato accentua l'importanza del raffreddamento delle soluzioni, mentre nella rapida sintesi di Wargo si dà maggiore importanza all'ebollizione, cioè all'effetto del diminuire della pressione negli strati più alti. È ovvio che ambedue i processi possono entrare in gioco; non mi meraviglierei se in futuro si dimostrasse che in alcuni giacimenti prevale l'ebollizione e in altri il raffreddamento.

Dal punto di vista della ricerca, gli studi svolti specialmente a Carlin hanno mostrato la cruciale importanza delle associazioni geochimiche. Disgraziatamente, composti di mercurio e d'arsenico non formano un alone regolare attorno al giacimento e si trovano in quantità molto variabili. Bisogna ricordare a questo proposito che le dimensioni spaziali dei giacimenti scoperti sono piccole: Carlin copre un'area inferiore a 1,5 chilometri quadrati; anche se si scoprissero in futuro giacimenti di maggiore estensione, per esempio di 10 chilometri quadrati, a scala geologica, l'obiettivo sarebbe sempre difficile per la sola esplorazione geochimica.

Tentiamo ora di indicare i metodi di ricerca che saranno adottati nel prossimo futuro. Per un lungo periodo, le ricerche verranno continuate nei distretti minerari che già hanno avuto produzione d'oro anche da piccoli giacimenti; in queste aree, le probabilità di scoprire nuovi giacimenti di tipo tradizionale (placer, Mother Lode, Homestake, Witwatersrand) sono minime, almeno nei paesi più popolati. La ricerca, quindi, punterà verso i possibili giacimenti di oro invisibile, simili, anche se non uguali, a quelli di Carlin, ma con caratteristiche geologiche differenti. Il primo passo nell'esplorazione continuerà a essere quello geologico, che può indicare le aree dove possono essere avvenute intrusioni magmatiche nel passato prossimo o remoto, o anche in tempi recenti. La geologia può indicare la natura geologica delle rocce nelle quali le masse plutoniche si sono intruse.

Le informazioni geologiche di maggior interesse comprendono: l'età e la petrografia della massa intrusa; l'età e la petrografia delle rocce entro cui le rocce plutoniche si sono intruse; le caratteristiche tettoniche contemporanee all'intrusione; le caratteristiche tettoniche attuali; la permeabilità delle rocce ospiti, tanto nel passato quanto attualmente; la profondità del giacimento che si spera di trovare; infine lo studio sistematico delle manifestazioni geotermiche, attive e fossili.

È chiaro che la geologia non può raggiungere i suoi obiettivi con la sola osservazione delle rocce affioranti in superficie, estrapolando poi in profondità. Anche nell'esplorazione aurifera, come in quasi tutti i settori dell'esplorazione mineraria moderna, il geologo e il geofisico devono procedere nella più stretta collaborazione, tanto che non è possibile delimitare i confini tra i due metodi di indagine. È vero che per il momento i metodi geofisici sono poco o punto utilizzati nell'esplorazione aurifera; ma nella nuova fase di ricerca, che ha per obiettivo principale la scoperta di giacimenti d'oro invisibile, la geofisica assumerà importanza determinante. Già i raffinati metodi geofisici di uso corrente possono essere utilizzati per ottenere informazioni preziose sulle intrusioni antiche, recenti e in atto. Le indagini geochimiche già forniscono un validissimo aiuto nella comprensione dei fenomeni geotermici, che sono certamente legati alla genesi dei giacimenti tipo Carlin. Il campo d'azione delle tecniche geofisiche attuali risulta invece limitato dalla piccola dimensione dei giacimenti auriferi, che sono anomali a scala geologica. Bisognerà quindi stimolare l'immaginazione dei geofisici, così flessibile e feconda, verso lo sviluppo di nuove tecnologie strumentali e teoriche.

Da pochi decenni, tutti i settori dell'esplorazione mineraria vanno trasformandosi rapidamente sotto l'impulso della rivoluzione geochimica, che in un certo senso ricorda la rivoluzione geofisica iniziata subito dopo la prima guerra mondiale.

I giacimenti d'oro, come abbiamo già detto, sono anomalie geochimiche molto piccole; anche le sostanze che accompagnano l'oro, come l'arsenico e il mercurio, già comunemente utilizzate nella prospezione e nella valutazione dell'oro, formano attorno ai giacimenti aloni mal definiti e possono servire soltanto nella fase preliminare della ricerca. La pirite, specie minerale associata all'oro nei giacimenti tipo Carlin, è troppo comune e diffusa e dà soltanto indicazioni molto generiche. Si sente il bisogno di nuovi concetti; Helgeson pensa che vi si possa giungere per via teorica, specialmente studiando gli aspetti termodinamici che guidano la soluzione, la mobilizzazione e la precipitazione dell'oro e dei minerali che l'accompagnano nei giacimenti di oro invisibile. Wargo pensa che la comprensione della storia termale dei giacimenti progredirà specialmente grazie allo studio geochimico delle miniere; da un punto di vista più pragmatico, richiama l'attenzione sugli aloni di calcite e di silice. Come si vede, Helgeson e Wargo tendono agli stessi scopi lungo vie confluenti.

Schizzo di Hansen e Kerr (*modificato*) di parte del grande scavo che permette di utilizzare il minerale d'oro della miniera di Carlin a cielo aperto. Lo scavo raggiunge una profondità di circa 180 metri. Sulla parete, si vede la faglia di sovrascorrimento delle Roberts Mountains (*linea tratteggiata con triangolini in colore*). Si noti la ripetizione delle serie sopra e sotto la faglia. I triangoli indicano la direzione della faglia, che ha portato il Siluriano (*Sr*) al di sotto delle rocce più antiche dell'Ordoviciano (*Ov*: formazione Vinini). La doppia linea tratteggiata indica la faglia di nord-est.

LOCALITÀ	ORO	ARGENTO
CHAMPAGNE POOL, WAIOTAPI	8	175
BOTOKAWA (POZZO)	70	30
OHAKI POOL, BROADLANDS	85	500
BROADLANDS (POZZO)	55	200

Concentrazioni di oro e argento in alcuni sinter della Nuova Zelanda, in parti per milione.

Poiché sono giunto alla fine di questo articolo, mi sembra doveroso aggiungere qualche ovvia osservazione sulle possibilità di ricerca d'oro invisibile in Italia. Faccio questo con estrema riluttanza, perché non sono aggiornato sulla situazione della ricerca in Italia e quindi corro il rischio di portare nottole ad Atene. Inoltre, non conosco di prima mano la situazione geologica delle vecchie miniere e ben poco sono riuscito a racimolare dalla letteratura. Invito quindi il lettore a leggere con estremo senso critico quel poco che dirò. Il mio unico scopo è quello di stimolare l'immaginazione esplorativa dei colleghi meglio informati di me. Malgrado le mie informazioni limitate, ricorderò alcuni fatti che mi lasciano perplesso.

Sappiamo che nell'antichità l'Italia era considerata un paese con alta produzione d'oro: per molti secoli - probabilmente dall'VIII secolo avanti Cristo fino al massimo splendore dell'Impero Romano - gli schiavi estrassero l'oro dalle sabbie dei fiumi alpini e da qualche filone di quarzo. Abbiamo già ricordato la miniera di Pestarena, dalla quale si estraevano piccole quantità d'oro fino alla fine della seconda guerra mondiale. Anche ora, del resto, si estraggono ridotte quantità d'oro dalle sabbie del Ticino.

Queste attività non hanno più alcuna importanza industriale; dimostrano però che l'oro ha potuto sciogliersi, muoversi e concentrarsi in filoni; anche le piccole miniere d'oro della zona di Carlin erano esaurite da vari decenni, prima della scoperta dell'oro invisibile.

Un altro punto che merita attenzione è l'esistenza delle intrusioni granitiche mioceniche della Toscana; secondo Giorgio Marinelli dell'Università di Pisa, esistono altre intrusioni simili della stessa età o più recenti. Alcune di esse hanno dato origine alle mineralizzazioni di solfuri (mercurio del Monte Amiata, pirite di Boccheggiano). Alcuni dei grandi campi geotermici della Toscana e del Lazio hanno con ogni probabilità un'origine simile. Non so se siano stati fatti gli studi geochimici atti a rivelare l'esistenza dell'oro invisibile nelle rocce incassanti e nello stesso minerale sfruttato. Queste indagini vanno certamente consigliate e si deve ricordare che nella pratica oggi comunemente seguita si analizzano migliaia di campioni prima di giungere a una decisione negativa; le analisi, naturalmente, non devono essere limitate all'oro e ogni area richiede un programma geochimico *ad hoc*.

Consiglierei anche lo studio geochimico e petrografico delle carote e dei fluidi provenienti dai pozzi geotermici già perforati o in perforazione, comprendendo anche i campioni dei pozzi spia per la misura della temperatura. I programmi di questo tipo richiedono tempo, pazienza e laboratori convenientemente attrezzati.

Dal lato positivo, bisogna notare il notevole sviluppo delle ricerche geologiche verificatosi specialmente negli ultimi venti anni; l'alto livello raggiunto nel campo della petrografia e della vulcanologia, specialmente per merito della scuola fondata da G. Marinelli dell'Università di Pisa; il fiorire della geochimica, tanto teorica quanto applicata all'esplorazione mineraria, che si può dire introdotta e sviluppata in Italia per merito di Franco Tonani, dell'Università di Palermo. In conclusione, l'Italia ha geologi, petrografi, geochimici e geofisici di alto livello ed essi hanno a disposizione le attrezzature più moderne. La fase preliminare dell'esplorazione aurifera richiede soltanto modesti capitali, che presumo si possano facilmente reperire.

Gli anelli della corrente del Golfo

La corrente del Golfo e altre correnti danno origine a vortici che, catturando masse d'acqua con proprietà fisiche e biologiche diverse da quelle del mare circostante, creano ecosistemi mobili particolari

di Peter H. Wiebe

L'opinione che vaste aree oceaniche siano uniformi nelle loro proprietà fisiche, chimiche e biologiche è ben radicata. Nell'oceano Atlantico, per esempio, si è per lungo tempo ritenuto che il Mar dei Sargassi fosse essenzialmente un bacino di acqua stagnante con un profilo termico uniforme in funzione della profondità marina e con un'uniforme scarsità di forme di vita. Tale opinione sta oggi perdendo terreno. Nuovi metodi di ricerca che utilizzano flottiglie di navi, strumenti installati su aerei e telemisurazioni effettuate da aerei e da satelliti rivelano schemi di variazione su scale diverse di spazio e di tempo insospettati fino a vent'anni fa. Ne emerge che gli stessi processi su grande scala della corrente oceanica e del clima mondiale, i quali favoriscono e mantengono la distribuzione della flora e della fauna marine su scale di migliaia di chilometri, perturbano però questi sistemi ecologici con la formazione di vortici dell'ordine di centinaia di chilometri.

I vortici più spettacolari sono anelli chiusi, anche di 300 chilometri di diametro, che si staccano dalle correnti come avviene per la corrente del Golfo, che scorre attraverso l'Atlantico a nord del Mar dei Sargassi. Quando un anello si forma, una colonna d'acqua si sposta da un fianco della corrente all'altro dove vi è un ambiente con condizioni di temperatura, salinità, contenuto in ossigeno e distribuzione delle forme di vita completamente differenti. Numerosi anelli possono trovarsi contemporaneamente nello stesso bacino oceanico ed è questo il motivo per cui il Mar dei Sargassi e altre regioni oceaniche sono dei veri e propri mosaici.

Il mio interesse personale per i vortici oceanici venne stimolato poco dopo essermi diplomato ed essere entrato a far parte, nel 1969, del personale della Woods Hole Oceanographic Institution. Udii le descrizioni di Peter Saunders, della Woods Hole, sui suoi tentativi di seguire l'evoluzione di un vasto vortice protetto dalla corrente del Golfo e immediatamente intuii che la persistenza di un tale vortice (essi possono durare per mesi, persino per un anno o due) deve avere conseguenze biologiche. Notai anche che nessuno aveva ancora cercato di studiare le relazioni tra i vortici e le varie forme di vita. Occorsero altri tre anni, per me e per i miei colleghi, per imbatterci in un vortice e iniziare a studiare la sua biologia.

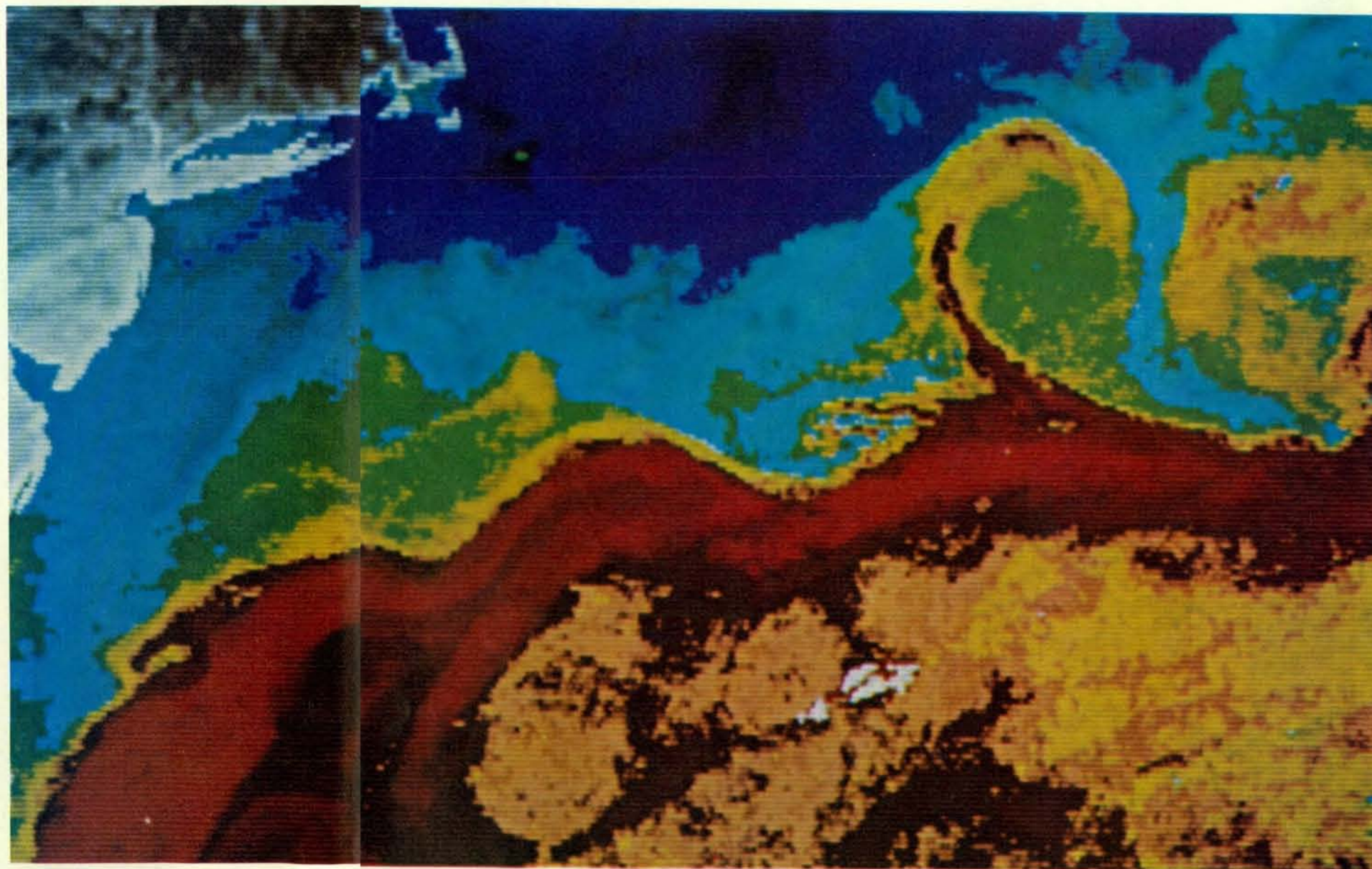
Quattro anelli occupano l'immagine a destra che mostra in falsi colori la distribuzione dell'intensità della radiazione infrarossa emessa dalla superficie dell'oceano Atlantico nordoccidentale tra il 12 e il 15 settembre 1981. Questo ci consente di dedurre la temperatura superficiale dell'acqua. Due anelli si trovano a sud della corrente del Golfo, la corrente rossa che serpeggia attraverso l'immagine da sud-ovest a nord-est. Essi sono le due macchie ovali color bronzo a ovest del piccolo ammasso bianco di nubi. In questo caso una corona circolare d'acqua staccatasi dalla corrente del Golfo circonda un nucleo di acqua fredda asportata dalla Slope Water (acqua di scarpata) che appare in varie sfumature di blu. Nel contempo due anelli si trovano a nord della corrente del Golfo: uno, proprio a sud di Capo Cod, ha un nucleo verde parzialmente circondato di giallo, l'altro, che appare più completo, si trova verso il margine orientale dell'immagine. Si è formato due mesi prima che l'immagine venisse ripresa, ma la sua rotazione in senso orario sta ora trascinando una scia rossa e gialla di acqua della corrente del Golfo che lo circonda per tre quarti. In questi anelli una corona circolare di corrente del Golfo circonda un nucleo di acqua calda che la corrente ha prelevato dal Mar dei Sargassi a sud. I dati da cui è stata ottenuta l'immagine sono stati registrati da un sensore all'infrarosso a bordo del satellite NOAA-7 e sono stati elaborati con l'aiuto di un calcolatore da Robert H. Evans, Otis B. Brown e James W. Brown dell'Università di Miami.

Il termine «anello», riferito a vortici chiusi che si staccano dalla corrente del Golfo, deriva da una proposta di Frederick C. Fuglister della Woods Hole, un pioniere nello studio della loro fisica, che se ne occupò dalla fine degli anni cinquanta. Il termine nasce dal fatto che un anello di acqua, separatosi dalla corrente del Golfo, scorre intorno a un nucleo d'acqua nettamente distinto tanto dalla corrente anulare quanto dall'acqua in cui l'anello, con relativo nucleo, è inglobato. Gli oceanografi avevano incontrato questo tipo di anelli durante crociere a sud della corrente del Golfo negli anni trenta e quaranta, ma fu solo in occasione dell'operazione Cabot, un rilevamento compiuto sulla corrente del Golfo intrapreso nel 1950 con sei navi dalla Woods Hole, insieme ad altre istituzioni, che un anello venne osservato nel momento della formazione.

Il processo di formazione si instaura quando la corrente del Golfo comincia ad assumere un andamento a meandri mentre scorre verso nord-est allontanandosi dal continente nordamericano dopo aver doppiato il Capo Hatteras nella North Carolina. Se il meandro diventa un cappio, si staccherà dalla corrente, come avviene per un'ansa fluviale che, strozzandosi, si trasforma in lanca o laghetto anulare. Il risultato sarà un anello di acqua della corrente del Golfo, rotante velocemente attorno a un nucleo di acqua marina di diversa origine. Un anello che si forma a nord della corrente avrà un nucleo d'acqua catturata da sud, cioè dal Mar dei Sargassi e quest'acqua sarà calda. Il Mar dei Sargassi è infatti noto per avere un esteso strato in cui la temperatura dell'acqua è di 18 gradi centigradi più o meno un grado. Tipicamente, questo strato comincia a una profondità di circa 200 metri e termina a circa 600 metri. Un anello che si forma a sud della corrente del Golfo avrà, invece, un nucleo d'acqua catturato da nord, e precisamente dal regime di acqua di scarpata *(Slope Water)* che si estende da Capo Hatteras fino ai Grand Banks al largo di Terranova e giace tra la corrente del Golfo e la piattaforma continentale dell'America Settentrionale. Quest'acqua è fredda; ha infatti una struttura caratteristica tale che la temperatura a una data profondità è inferiore di una decina di gradi a quella dell'acqua del Mar dei Sargassi.

Gli anelli neoformati sono vasti. Quelli a nucleo freddo hanno generalmente un diametro che va da 150 a 300 chilometri, ossia hanno un'area superficiale pari a tre volte quella della Sardegna. Inoltre, essi si estendono fino al fondo marino, che giace a una profondità compresa tra 4000 e 5000 metri. Gli anelli a nucleo caldo sono in genere un po' più piccoli: hanno un diametro compreso tra 100 e 200 chilometri e la loro estensione verticale è spesso avvertibile solo fino a 1500 metri di profondità. Essi però sono stati studiati molto meno di quelli a nucleo freddo e ulteriori misurazioni potranno cambiare le nostre attuali conoscenze in proposito.

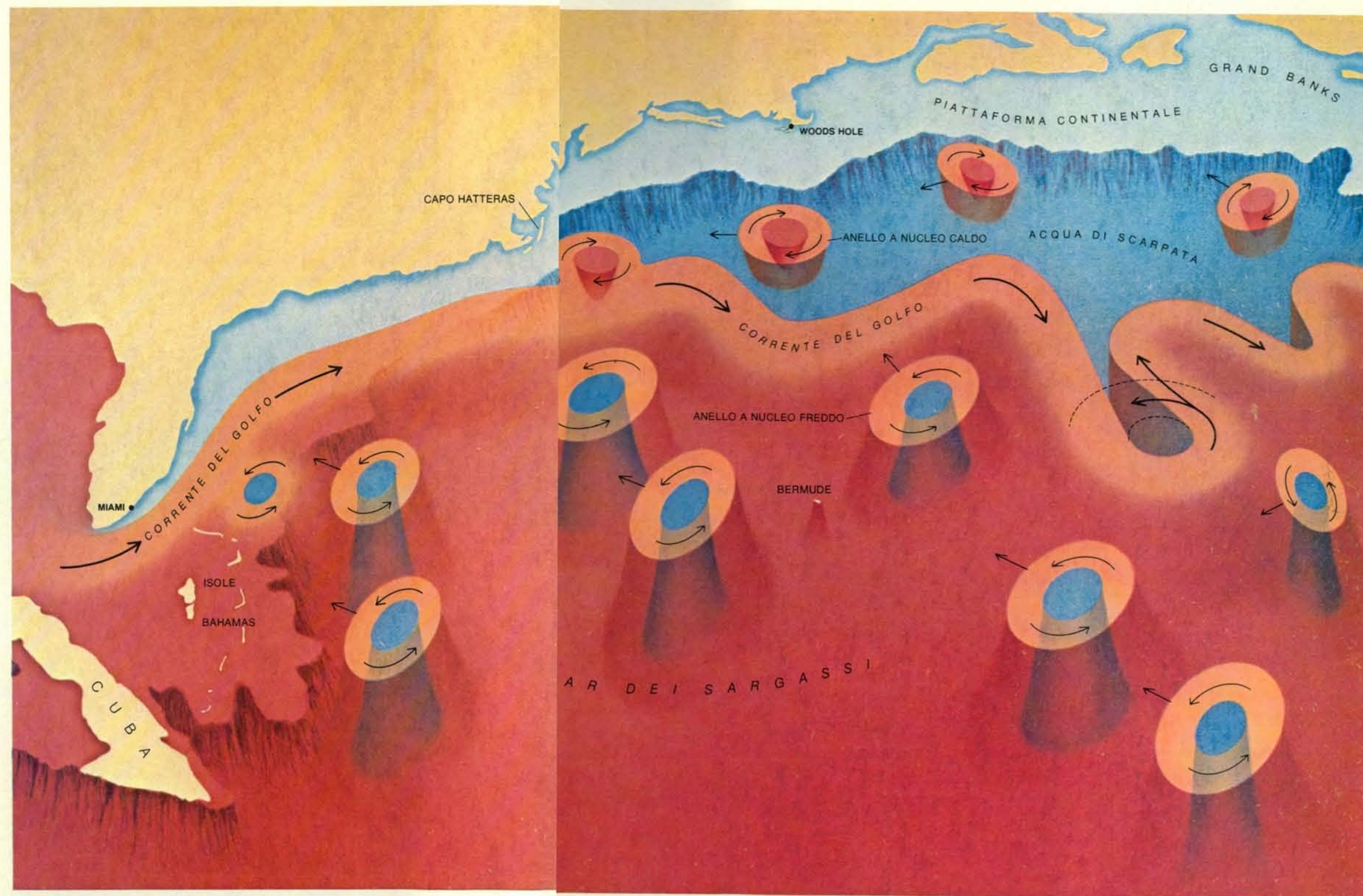
Ogni anello, a nucleo caldo o a nucleo freddo che sia, comprende una porzione della corrente del Golfo che può raggiun-

gere un'ampiezza di 100 chilometri. La velocità di rotazione dell'acqua di questa corona circolare può anche raggiungere i tre nodi, ovvero 1,5 metri per secondo, a una distanza di 30-40 chilometri dal centro dell'anello. Boe equipaggiate con un radiotrasmettitore i cui segnali sono captati da un satellite compiono il giro di un anello giovane in un periodo compreso tra i due e cinque giorni. La velocità di rotazione decresce verso il centro dell'anello e anche all'aumentare della distanza nell'acqua che circonda l'anello, nonché con la profondità. Ciononostante, velocità maggiori di 0,5 metri per secondo sono state misurate a profondità di 400-500 metri al di sotto delle forti correnti di superficie.

È molto facile rendersi conto delle condizioni anomale dovute alla presenza di un anello a nucleo freddo nel Mar dei Sargassi osservando sulla carta l'andamento verticale della temperatura nel mare e, in particolare, l'isoterma di 15 gradi, ossia la curva che indica la profondità dell'acqua la cui temperatura è di 15 gradi. Nel Mar dei Sargassi l'isoterma di 15 gradi si trova tipicamente a una profondità compresa tra i 600 e i 700 metri dalla superficie. Presso il margine meridionale della corrente del Golfo essa sale gradualmente, ma nella corrente stessa la sua salita è molto repentina. Infatti, il gradiente termico attraverso la corrente del Golfo è uno dei più ripidi rispetto a tutti gli oceani del mondo. Tant'è che il margine settentrionale della corrente è chiamato Muro Freddo. Lungo il Muro Freddo l'isoterma di 15 gradi si trova a una profondità di 100 metri e nell'acqua di scarpata, a nord della corrente, si trova ancora più vicina alla superficie.

La struttura termica di un anello a nucleo freddo appena formato mostra gradienti ripidi analoghi. Consideriamo l'anello a nucleo freddo denominato Bob, che si formò nel febbraio 1977 e venne studiato durante tutto l'anno, dalla sua formazione alla sua estinzione. Al centro dell'anello le isoterme delle temperature comprese tra i 10 e i 16 gradi centigradi superavano di 600 metri le loro normali profondità del Mar dei Sargassi. La densità dell'acqua marina, il cui andamento è essenzialmente parallelo a quello della temperatura, mostrava aumenti analoghi.

Il maggior spessore dell'acqua marina densa significava che l'anello Bob inglobava un immenso serbatoio di energia potenziale. Significava anche che l'anello Bob esercitava una pressione sull'acqua circostante, pressione che era aumentata dalla forza centrifuga esercitata dalla rotazione dell'anello. Inoltre, essa era modulata poiché la superficie del mare in un anello devia dall'essere gravitazionalmente livellata. (Misurazioni effettuate dai satelliti indicano che la superficie al centro di un anello è depressa di circa 0,5 metri.) La pressione verso l'esterno esercitata da un anello è ampiamente bilanciata dall'effetto Coriolis, una forza che agisce sugli anelli quando migrano nell'oceano e che insorge a causa della rotazione della Terra. Più generalmente, si crede che la migrazione di un anello e la sua eventuale scomparsa siano soprattutto dovute agli squilibri tra pressione, forza centrifuga e forza di Coriolis. Altre forze, come l'attrito, sono soltanto l'un per cento delle forze primarie anzidette.

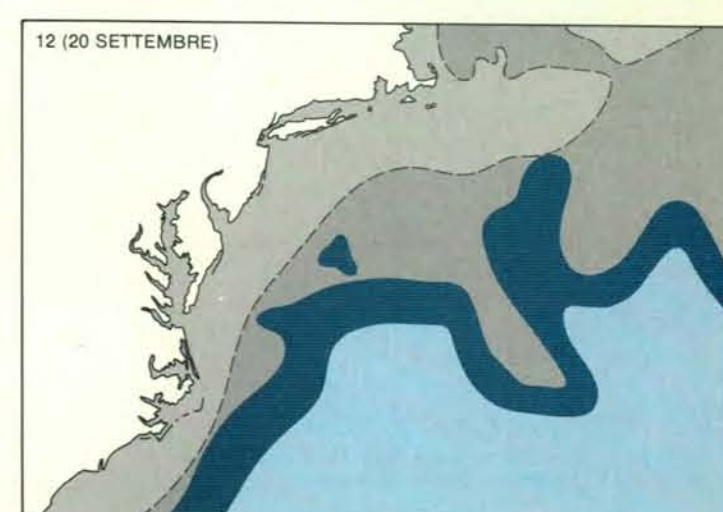
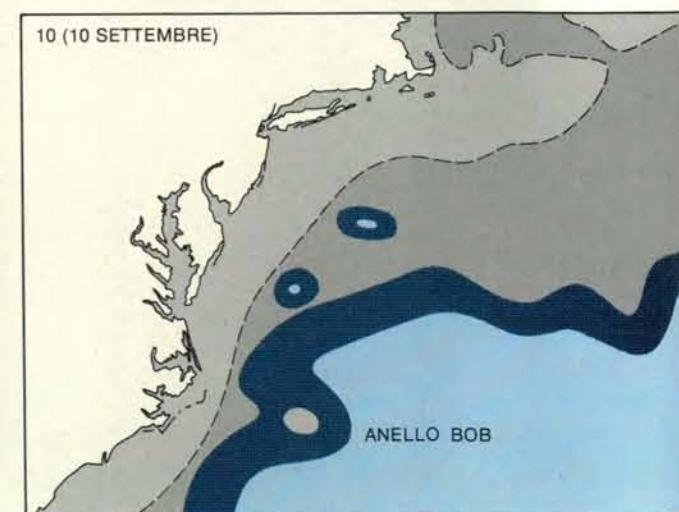
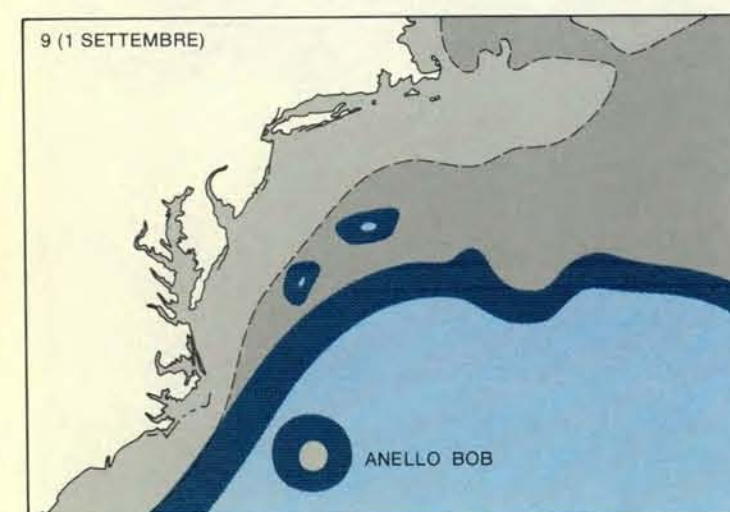
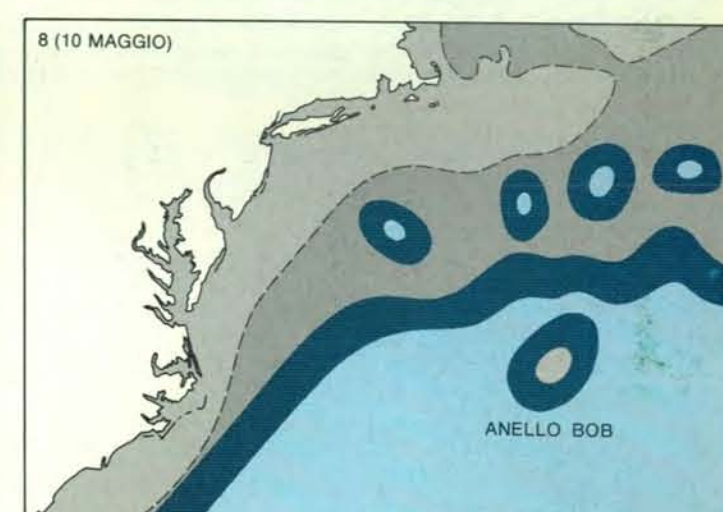
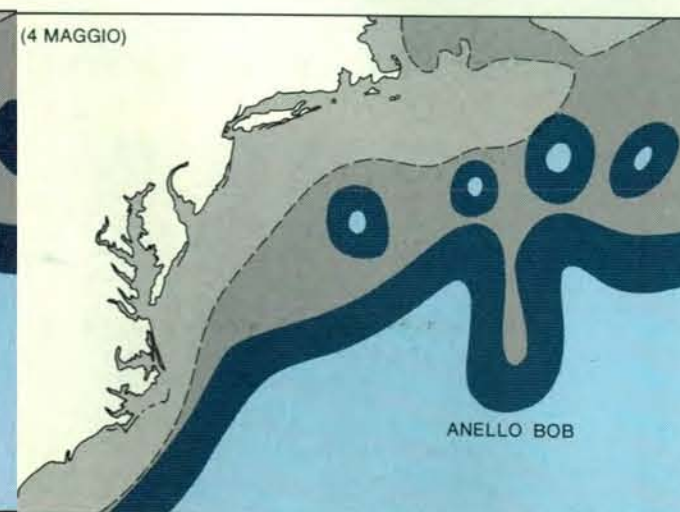
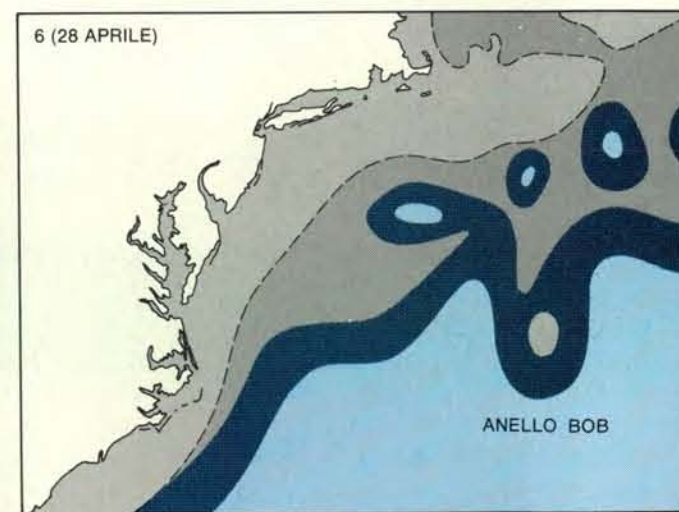
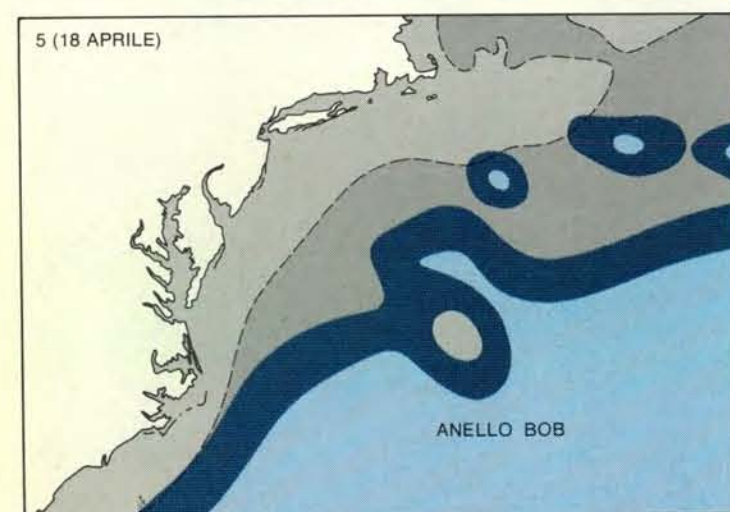
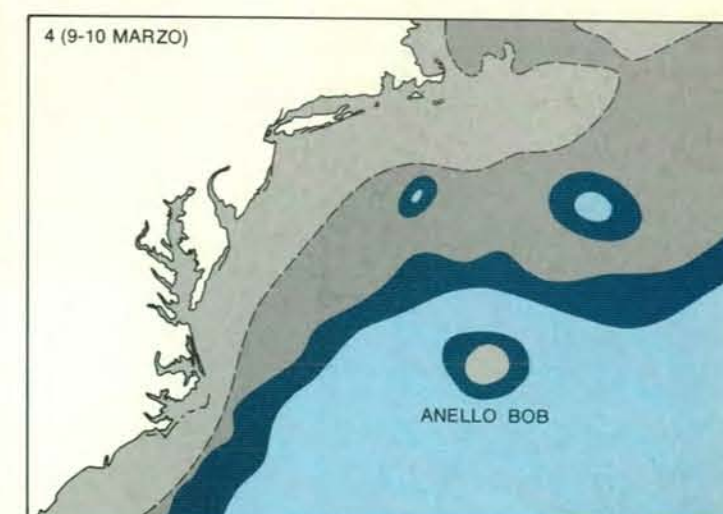
La distribuzione della salinità fornisce un ulteriore chiarificante indizio sulla struttura degli anelli. In particolare, in un anello a nucleo freddo, la salinità nelle prime centinaia di metri dalla superficie può essere di una parte per mille più bassa della salinità dell'acqua nel circostante Mar dei Sargassi alla stessa temperatura (ma a maggiore profondità). Tale anomalia, del tutto sorprendente per l'acqua del mare aperto, può persistere per 18 mesi e anche oltre; dimostra pertanto che una parte sostanziale di un anello rimane immutata per un certo periodo di tempo. Dimostra inoltre che un anello è completamente differente da una formazione ondosa in cui acqua fredda si porti temporaneamente in superficie nell'oceano.

La distribuzione dell'ossigeno disciolto nell'oceano ha pure valore diagnostico. L'acqua della corrente del Golfo proveniente dai Caraibi ha una bassa concentrazione di ossigeno per temperature comprese tra otto e 18 gradi centigradi; l'acqua di scarpata e i nuclei degli anelli giovani a nucleo freddo hanno una concentrazione elevata; l'acqua del Mar dei Sargassi e i nuclei degli anelli giovani a

La tipica distribuzione degli anelli è illustrata in un disegno basato sull'analisi della temperatura oceanica e su altri dati da Philip L. Richardson della Woods Hole Oceanographic Institution. In generale gli anelli a nucleo freddo, che occupano il Mar dei Sargassi, hanno un diametro che può raggiungere i 300 chilometri e possono estendersi fino al fondo marino a una profondità compresa tra 4000 e 5000 metri. Gli anelli a nucleo caldo, che interessano l'acqua di scarpata, sono più piccoli e hanno uno spessore inferiore. Gli anelli si formano nella parte orientale della regione illustrata e si muovono in modo erratico verso sud-ovest indebolendosi in capo a parecchi mesi. In ogni istante gli anelli a nucleo caldo possono avere un'estensione pari al 40 per cento dell'acqua di scarpata e gli anelli a nucleo freddo pari al 15 per cento del Mar dei Sargassi. Nel disegno la scala verticale dell'oceano è stata esagerata di circa un fattore 100.

nucleo caldo hanno una concentrazione intermedia. Da queste differenze si vede come misurazioni molto accurate della concentrazione di ossigeno lungo i fianchi degli anelli rendano possibile classificare l'acqua marina a seconda della sua origine. Le misurazioni, effettuate da Dana R. Kester dell'Università di Rhode Island, mostrano chiaramente che l'acqua dei fianchi di un anello a nucleo freddo e quella del circostante Mar dei Sargassi sono intercalate. Il tasso di scambio di queste intercalazioni non è stato ancora determinato. Infatti, non è ancora chiaro se esse rappresentino scambi d'acqua o se, invece, siano delle strutture che persistono dal momento in cui l'anello si è formato.

Quanti anelli si formano in un anno? Quanto tempo durano? Dove vanno e come si evolvono?

Fuglister ha valutato che il numero di anelli che si formano su ogni lato della corrente del Golfo in un anno varia da cinque a otto. Il suo metodo, indiretto, basava i calcoli su due fattori. Primo, la differenza in volume tra l'afflusso di acqua della corrente del Golfo a Capo Hatteras e il suo deflusso ai Grand Banks nel corso di un anno; secondo, la variazione dell'area superficiale dell'acqua di scarpata causata dalle deviazioni della corrente del Golfo dalla sua posizione media.

Un censimento di anelli a nucleo caldo, effettuato da J. L. Chamberlin e dai suoi

La storia di un anello è esemplificata da quella dell'anello a nucleo freddo Bob, nato nel febbraio del 1977. Dapprima la corrente del Golfo formò un meandro rivolto a sud *(1)*, divenuto poi un cappio *(2)*, il quale si chiuse *(3)* e si staccò dalla corrente *(4)*, divenendo un anello. Cinque settimane dopo si ricongiunse alla corrente, spostandosi dal ramo occidentale a quello orientale di un altro meandro in formazione,

che in seguito divenne un anello a nucleo caldo *(5, 6)*. In maggio si staccò nuovamente *(7,8)*, restando stazionario per alcune settimane. In settembre, però, si avvicinò di nuovo alla corrente del Golfo *(9)*. Qui si fuse nella corrente e scomparve *(10, 11, 12)*. La storia è stata ricostruita da Richardson sulla base di immagini riprese nell'infrarosso. Le forme discoidali a nord della corrente del Golfo sono anelli a nucleo caldo.

colleghi del National Marine Fisheries Service di Narragansett, Rhode Island, è in accordo con la stima di Fuglister per quanto riguarda il lato settentrionale della corrente. Il censimento ha registrato da sette a nove anelli a nucleo caldo per anno dal 1975 al 1978. Talvolta durante l'anno nessun anello era presente, talaltra esistevano sei o sette anelli. Quando molti anelli erano presenti, essi occupavano anche il 40 per cento dell'area superficiale dell'acqua di scarpata. Nessun censimento analogo è stato tentato a sud della corrente, in parte perché il costo dei rilevamenti su vasta scala fatti da navi e da aerei è elevato. Inoltre, la sorveglianza degli anelli dai satelliti è compromessa per gran parte del tempo dalla copertura di nubi. (Le osservazioni dai satelliti sono effettuate alle lunghezze d'onda infrarosse intensamente emesse da corpi caldi. Le immagini risultanti mostrano pertanto lo schema di distribuzione della temperatura alla superficie del mare, incluso lo schema caratteristico degli anelli.)

Un conteggio diretto degli anelli a nucleo freddo attende l'avvento di sensori a microonde di cui si prevede l'installazione a bordo dei satelliti nel corso di questo decennio. Questi sensori rileveranno le radiazioni dell'oceano, e quindi la sua temperatura, a lunghezze d'onda maggiormente in grado di penetrare attraverso le nubi. Essi misureranno anche la profondità del mare, un'ulteriore indicazione della presenza di anelli. Per ora Philip L. Richardson, della Woods Hole, ha analizzato le registrazioni della distribuzione della temperatura nel Mar dei Sargassi e anche le registrazioni di un rilevamento del Mar dei Sargassi, fatto con navi e satelliti nella primavera del 1975. Egli ne deduce che in ogni momento sono presenti circa dieci anelli a nucleo freddo.

Gli anelli più giovani e più netti si trovano in una regione in cui gli anelli si formano di preferenza proprio a sud della corrente del Golfo tra 75 e 60 gradi di longitudine ovest. Anelli più vecchi si trovano a sud e a ovest di questa regione. Nel Mar dei Sargassi, a sud di 30 gradi di latitudine nord e a est di 65 gradi di longitudine ovest, non è mai stato rilevato alcun anello. Questa regione, però, non è stata bene esplorata. Tutto sommato sembra lecito affermare che gli anelli a nucleo freddo occupino un'area superficiale del Mar dei Sargassi compresa tra il 10 e il 15 per cento dell'area complessiva.

Il numero di anelli in un dato momento è ovviamente in parte determinato dal loro tempo di persistenza. Anche questo aspetto della storia degli anelli attualmente può essere trattato solo in modo limitato. Richardson e i suoi colleghi hanno piazzato delle boe in molti anelli e le hanno seguite tramite satellite per otto mesi. Nessuna di queste osservazioni, però, ha riguardato l'intera vita di un anello. Si è visto che alcuni anelli, identificati in successive immagini all'infrarosso, erano durati per due anni. Inoltre, un calcolo basato sulla quantità di energia potenziale, contenuta in un anello giovane, e una stima del tasso al quale tale energia viene dissipata fanno pensare a una durata di vita compresa tra due e quattro anni. Il fondersi dell'anello entro la corrente del Golfo ne abbrevia spesso la vita e, quindi, Richardson stima che la vita media di un anello a nucleo freddo vada da 12 a 18 mesi. Il massimo probabile è di tre anni.

Gli anelli si spostano in modo complesso. Oltre al moto vorticoso, essi migrano attraverso l'oceano; in effetti, il movimento di traslazione dell'acqua che compone un anello è ipotizzato fino a profondità di 1500-2000 metri in un modello ideato da Glenn R. Flierl del Massachusetts Institute of Technology. Tanto gli anelli a nucleo caldo quanto quelli a nucleo freddo, una volta staccatisi dalla corrente del Golfo, tendono a spostarsi verso ovest o sud-ovest a una velocità media dell'ordine di cinque centimetri per secondo (quattro chilometri al giorno). La velocità di traslazione varia: tipicamente esistono periodi in cui un anello si sposta a velocità superiore a 10 centimetri al secondo e altri periodi in cui è fermo.

Alcuni anelli si mantengono lontani dalla corrente del Golfo mentre si muovono verso sud-ovest lungo un «viale dei vortici», fintanto che arrivano in una regione a nord delle Bahamas e a est della Florida. Si tratta del «cimitero degli anelli», una zona in cui gli anelli si fondono con la corrente del Golfo e scompaiono. Sembra però che la maggior parte degli anelli interagisca ripetutamente con la corrente del Golfo. Gli anelli entrano in contatto con la corrente che li respinge verso nord-est a una velocità che raggiunge anche i 25 centimetri al secondo. Inoltre, se la fusione non è completa, essi si staccano, talvolta fortemente modificati. In alcuni casi risultano rimpiccioliti, ma dotati di maggior energia. Ognuno di questi cicli dura da uno a tre mesi. A causa delle numerose interazioni con la corrente del Golfo, la traiettoria che un anello a nucleo freddo compie attraverso il Mar dei Sargassi può descrivere una serie di ampi cappi in senso orario.

Nei periodi in cui un anello rimane isolato dalla corrente del Golfo, gradualmente decade. Per un anello a nucleo freddo il processo può essere evidenziato dal riscaldamento dell'acqua nel nucleo, dall'aumento della salinità in quest'acqua e dall'aumento della profondità della zona in cui il contenuto in ossigeno dell'acqua è minimo. (Nel Mar dei Sargassi settentrionale la zona in cui l'ossigeno è minimo si trova tra 700 e 1000 metri. Nell'acqua di scarpata, che costituisce il nucleo di un anello giovane a nucleo freddo, detta zona giace tra 100 e 200 metri.) L'indice migliore preso singolarmente del decadimento fisico di un anello a nucleo freddo è la velocità con cui aumenta la profondità delle isoterme. Ripetute misurazioni della distribuzione verticale della temperatura in numerosi anelli a nucleo freddo fanno ritenere che questa velocità sia compresa tra 0,3 e 0,4 metri al giorno. In generale i valori più bassi sono misurati negli anelli più vecchi. Le isoterme prendono la forma di una cupola di acqua fredda che collassa. Stanno anche a indicare la diminuzione dell'energia cinetica dell'anello e del volume d'acqua che l'anello trasporta.

La degradazione fisica di un anello è accompagnata da una profonda trasformazione della sua biologia. La ragione è semplice. Le forme di vita presenti nell'acqua fredda di scarpata sono del tutto differenti da quelle delle acque più calde del Mar dei Sargassi. In un anello giovane questa differenza si riscontra, mentre man mano che l'anello decade l'ambiente che favorisce l'esistenza delle forme di vita nel suo interno comincia a mutare.

Fondamentalmente nell'acqua del Mar dei Sargassi (così come nel nucleo di un anello giovane a nucleo caldo) vivono moltissime specie tropicali e subtropicali. L'entità delle popolazioni, però, è relativamente piccola e uniforme. L'acqua di scarpata, invece (come il nucleo di un anello giovane a nucleo freddo) è molto simile ai mari subartici situati più a nord. Essa sostenta un numero relativamente piccolo di specie, di cui alcune sono dominanti. La sua biomassa, ovvero il contenuto totale di materia vivente, è molto maggiore di quella delle acque più calde.

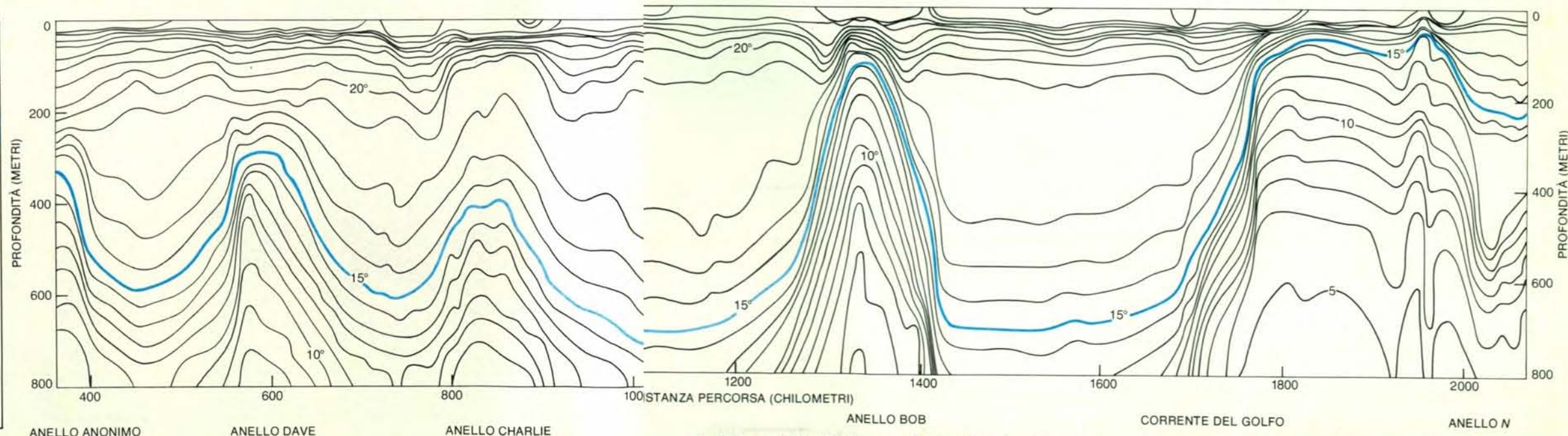
Mi concentrerò ora sulle forme planctoniche di vita marina. Il termine comprende sia vegetali (fitoplancton) sia animali (zooplancton) vaganti alla deriva in mare aperto. In generale il fitoplancton è costituito da vegetali unicellulari le cui dimensioni variano da cinque a 75 micrometri. Essi sono i produttori primari del mare dato che costruiscono molecole organiche da molecole inorganiche per mezzo della fotosintesi; essi, quindi, contengono clorofilla, la sostanza che media la fotosintesi. I costituenti dello zooplancton possono avere dimensioni alquanto maggiori rispetto a quelli del fitoplancton, ma generalmente non superano i quattro centimetri di lunghezza. Variano da organismi unicellulari, come i protozoi, a organismi più complessi, come i crostacei e le larve di molte specie di pesci. Vivono a spese del fitoplancton oppure a spese reciproche in una serie di relazioni altamente complesse.

Premesso tutto ciò, consideriamo nuovamente l'anello Bob. Quando con i miei colleghi lo studiai per la prima volta nell'aprile 1977, la distribuzione della clorofilla indicava notevoli differenze di concentrazione del fitoplancton tra il centro dell'anello e l'acqua circostante. Nel nucleo dell'anello, costituito da fredda acqua di scarpata, la clorofilla (e presumibilmente il fitoplancton) era altamente concentrata; misurammo 3,0 microgrammi di clorofilla per litro d'acqua. La distribuzione era circoscritta soprattutto ai 30 metri più superficiali del mare. Con la distanza dal centro la concentrazione diminuiva e si trovò che era massima tra i 50 e i 70 metri di profondità. Lo schema era tipico delle differenze tra il Mar dei Sargassi e l'acqua di scarpata.

La distribuzione dello zooplancton era alquanto più complessa. Ci aspettavamo di trovare concentrazioni elevate al centro dell'anello e una graduale diminuzione ai fianchi, cosicché la concentrazione in periferia uguagliasse quella tipica del Mar dei Sargassi. In effetti essa risultò elevata al centro. Nei resti della corrente del Golfo che circondava il nucleo, però, trovammo una zona in cui lo zooplancton era persino meno abbondante di quanto fosse nell'oceano circostante. Evidentemente la corrente del Golfo, al momento in cui l'anello Bob si era formato (cioè due mesi prima che vi arrivassimo), aveva una concentrazione di zooplancton inferiore a quella del Mar dei Sargassi e tale scarsità di forme di vita persisteva.

Le abbondanze globali relative al fitoplancton e allo zooplancton, nell'aprile 1977, riflettono le abbondanze di specie particolari. Per esempio, specie di acqua calda, come *Limacina inflata* (una chiocciola marina che vive nelle acque tropicali e subtropicali di tutti gli oceani della Terra), vennero trovate in abbondanza solo all'esterno del giovane anello Bob. Specie planctoniche, invece, che vivono nella fredda acqua di scarpata erano confinate nel nucleo dell'anello. Tali specie includevano *Limacina retroversa*, un'altra chiocciola marina, come pure *Nematoscelis megalops* (un piccolo crostaceo simile a un gamberetto) e *Pareuchaeta norvegica* (un piccolo crostaceo che manca di occhi composti). Le differenti distribuzioni verticali che misurammo per queste specie di acqua fredda nel nucleo dell'anello Bob, nell'aprile 1977, riflettevano le profondità alle quali dette specie vivono normalmente nell'acqua di scarpata.

Al nostro ritorno all'anello Bob, nell'agosto 1977, trovammo quanto avevamo già trovato in altri anelli a nucleo freddo

La struttura termica dell'Atlantico nordoccidentale è stata registrata lungo una rotta a zig-zag da Miami, in Florida, a Woods Hole, nel Massachusetts, tenuta dalla nave di ricerca *Endeavor* attraverso il nucleo di cinque anelli, compreso l'anello Bob, nell'agosto 1977. Nell'anello Bob, a nucleo freddo, l'isoterma di 15 gradi centigradi *(in colore)* aveva un livello di 600 metri più elevato rispetto a quello del Mar dei Sargassi circostante. Nell'anello N, invece, a nucleo caldo, l'isoterma era a un livello di circa 150 metri più basso del suo livello locale.

In un anello, la biologia si modifica man mano che l'anello invecchia dato che mutano le condizioni di temperatura, salinità e contenuto in ossigeno. Nel grafico viene indicata l'abbondanza di cinque piccole specie di crostacei, simili a gamberetti, nel nucleo dell'anello Bob e nel circostante Mar dei Sargassi, nell'aprile 1977 (quando l'anello aveva due mesi) e nell'agosto successivo. Nel corso di questi quattro mesi *Stylocheiron carinatum (a)*, una specie d'acqua calda che vive nei pressi della superficie, entrò nell'anello raggiungendone il centro. Nel contempo *S. suhmii (b)*, un'altra specie d'acqua calda che vive presso la superficie, entrò in minor misura. *S. elongatum (c)*, una specie d'acqua calda che vive molto al di sotto della superficie, non entrò quasi del tutto. *Euphausia brevis (d)*, una specie d'acqua calda che migra verso l'alto di notte e verso il basso di giorno, mostrò la stessa penetrazione nelle retate diurne e notturne. (Il colore indica la sua abbondanza diurna, mentre il nero indica quella notturna.) Per le altre specie le abbondanze diurne e notturne sono unite. *Nematoscelis megalops (e)* è una specie d'acqua fredda, rinvenuta solo all'interno dell'anello. Tra aprile e agosto si spostò più in profondità, cercando evidentemente le condizioni favorevoli alla sua sopravvivenza. La curva in colore di ogni grafico si riferisce all'isoterma di 15 gradi.

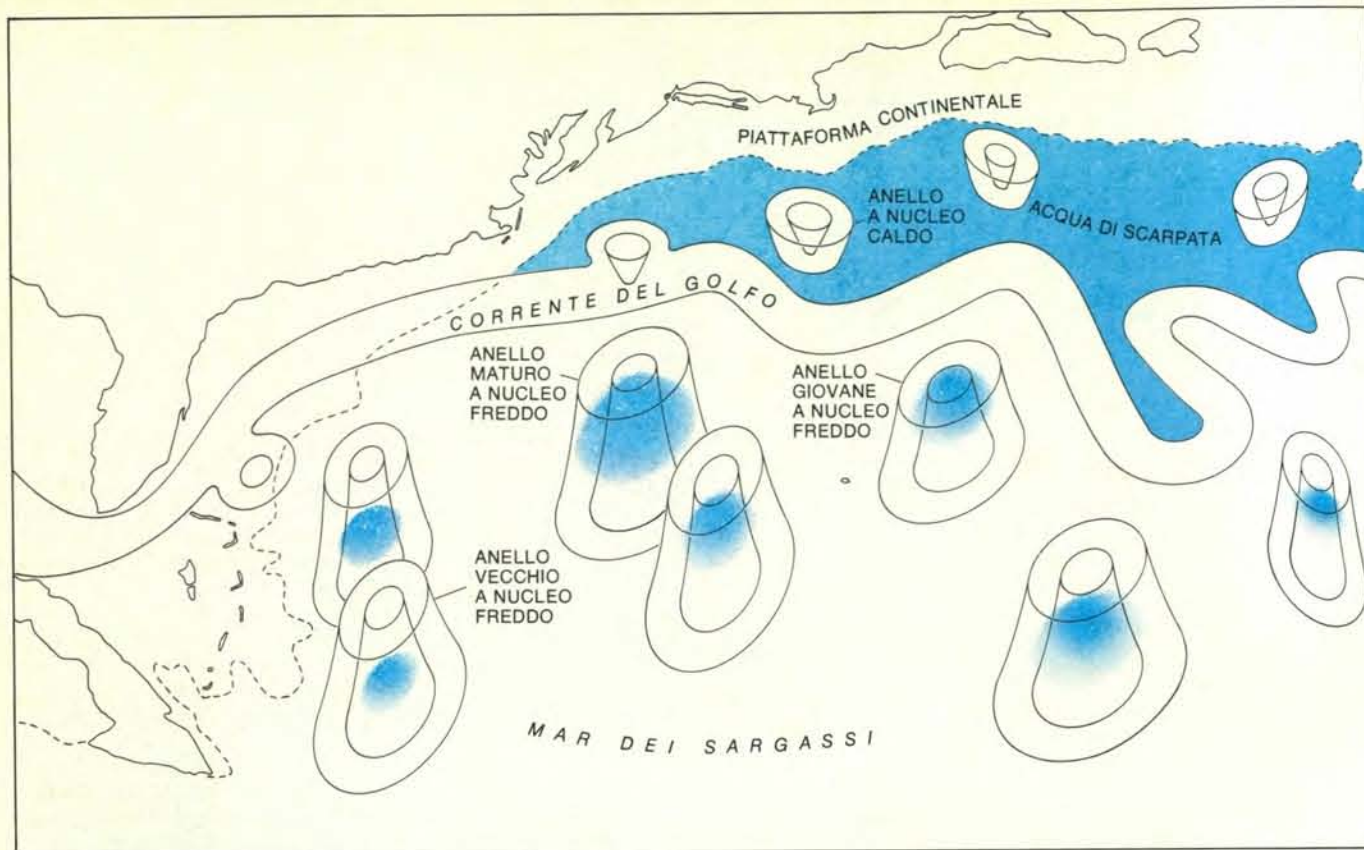
in fase di decadenza: le specie della comunità planctonica di acqua fredda dell'anello erano diminuite di numero. Se l'anello avesse continuato nel suo isolamento dalla corrente del Golfo, senza dubbio quelle specie sarebbero completamente scomparse dall'anello. I nostri studi dimostrano che il tempo necessario per l'estinzione varia da specie a specie: esso oscilla tra alcuni mesi e un anno e mezzo. I nostri studi dimostrano anche che la diminuzione del plancton in un nucleo freddo comporta una curiosa distribuzione delle forme di vita nella colonna d'acqua. Spesso i 200 metri più superficiali di un anello a nucleo freddo in fase di disfacimento, hanno significativamente meno biomassa dei 200 metri più superficiali del circostante Mar dei Sargassi. I successivi 800 metri sono significativamente più ricchi di quelli corrispondenti del Mar dei Sargassi. È già abbastanza significativo che la colonna d'acqua di un anello contenga un'abbondanza di zooplancton maggiore in quella del circostante Mar dei Sargassi per almeno un anno dalla sua formazione.

Sembra che i responsabili di questa curiosa distribuzione siano due processi simultanei. In primo luogo, il rimescolamento del mare a opera dei venti e il suo riscaldamento da parte del Sole influenzano la parte superiore delle acque. Di conseguenza l'affievolirsi in un anello a nucleo freddo delle condizioni che favoriscono le forme di vita caratteristiche dell'acqua fredda è più rapido in prossimità della superficie. Dal momento in cui un anello a nucleo freddo si forma, la parte superiore del suo nucleo può conservarsi, per un periodo da uno a tre mesi, poco diversa dall'acqua di scarpata.

In seguito ha inizio una sensibile trasformazione: diminuiscono tanto l'abbondanza di fitoplancton d'acqua fredda quanto le dimensioni degli organismi. Nell'anello Bob questo cambiamento fu sorprendente: tra aprile e agosto 1977 la concentrazione di clorofilla nel nucleo diminuì di un fattore otto. Non meraviglia il fatto che erbivori come *Limacina retroversa*, che vivono in prossimità della superficie, sembrino più sensibili alle modificazioni del loro ambiente di acqua fredda, in un anello a nucleo freddo maturo, rispetto a onnivori che vivono a maggiore profondità, come *Nematoscelis megalops* o carnivori come *Paraucheta norvegica.*

La seconda ragione della concentrazione di vita planctonica molto al disotto della superficie in un anello a nucleo freddo in disfacimento è costituita dal fatto che, a quanto pare, un certo numero di specie di acqua di scarpata migrano verso il basso nella colonna d'acqua nel tentativo di rimanere nel loro ambiente preferito. *Nematoscelis megalops* rende ben evidente questo spostamento. Esso è tipico dell'acqua di scarpata: nelle campionature eseguite nell'Atlantico nordoccidentale non l'abbiamo mai trovato a sud della corrente del Golfo, se non negli anelli a nucleo freddo. Nell'acqua di scarpata la maggior parte della popolazione di *N. megalops* si trova entro i 300 metri di profondità sia di giorno sia di notte.

Negli anelli giovani a nucleo freddo la distribuzione della popolazione è la stessa. Quando l'anello ha sei mesi di età, però, un numero considerevole di organismi vive al di sotto di 300 metri e organismi singoli vengono trovati fino a 800 metri di profondità. In anelli ancora più vecchi la maggior parte della popolazione vive al di sotto di 300 metri. L'aumento della profondità nella distribuzione verticale di *N. megalops* in un anello a nucleo freddo in decadenza coincide con l'aumento di profondità delle isoterme e anche con quello delle linee isoaline, che indicano la salinità dell'acqua. In realtà sembra che il 50 per cento centrale della popolazione adulta di *N. megalops* ricerchi una temperatura compresa tra otto e 12 gradi centigradi sia nell'acqua di scarpata sia in un anello.

Il movimento verso il basso di *N. megalops* significa che gli animali si allontanano dalla concentrazione di cibo più vicina alla superficie; quindi lo spostamento verso il basso coincide con un deterioramento delle condizioni fisiologiche e biochimiche della specie. Il declino fisiologico è dimostrato da una riduzione nel tasso di respirazione degli animali, quello biochimico da un aumento nel loro contenuto idrico e da una riduzione nel loro contenuto in lipidi, carbonio e azoto. Inoltre abbiamo trovato in anelli a nucleo freddo più vecchi una riduzione del tasso di crescita di esemplari di *N. megalops* rispetto al loro tasso di crescita nell'acqua di scarpata, e una riduzione del rapporto tra maschi e femmine. Concludendo, sembra che gli individui di *N. megalops*, in un anello a nucleo freddo in decadenza sof-

La distribuzione ipotetica di *Nematoscelis megalops*, indicata da sfumature di colore, riflette la propensione della specie per acque con temperatura di circa dieci gradi centrigradi. *N. megalops* è abbondante a nord della corrente del Golfo. A sud può vivere solo in anelli giovani. In anelli maturi migra a maggiori profondità. In anelli vecchi tende all'estinzione. Questa distribuzione mette in dubbio la precedente ipotesi secondo cui le condizioni sono uniformemente adatte o inadatte per la sopravvivenza di una data specie nell'ambito di ampie distese di mare.

frano la fame man mano che i livelli di cibo diventano inadeguati alla crescita e alla riproduzione. L'anello a nucleo freddo più vecchio che abbiamo campionato aveva 17 mesi. In esso non abbiamo trovato nessun esemplare di *N. megalops,* anche se precedentemente ce n'erano.

Scompaiono altre specie per mancanza di cibo? Sembra probabile; eppure per ogni specie planctonica esiste l'ulteriore possibilità che il moto e il disfacimento di un anello possano provocare l'espulsione di individui verso il circostante Mar dei Sargassi. Il modello progettato da Flierl al MIT chiaramente fa pensare che, quando un anello migra attraverso il Mar dei Sargassi, solo una parte della colonna d'acqua dell'anello venga in effetti trascinata. Secondo il modello, la parte intrappolata di un anello ha la massima estensione in superficie e va restringendosi con l'aumentare della profondità fino a scomparire tra 1000 e 1500 metri. Di conseguenza individui di una specie come *N. megalops,* che si spostano sempre più in profondità nella colonna d'acqua di un anello in decadenza, possono finire al di sotto della parte intrappolata dell'anello, che può abbandonarli dietro di sé. Le nostre analisi personali sulla distribuzione di specie di acqua fredda come *N. megalops* e *P. norvegica* sui fianchi di anelli a nucleo freddo indicano che alcuni esemplari vengono in realtà allontanati da processi fisici.

Mentre varie specie di acqua fredda cominciano a trovare, in un anello a nucleo freddo in decadenza, condizioni meno adatte alla loro sopravvivenza, alcune specie di acqua calda presenti nel circostante Mar dei Sargassi cominciano a trovarle favorevoli. Presso la superficie il fitoplancton d'acqua fredda cede il posto ad alcune forme più piccole e meno abbondanti tipiche del Mar dei Sargassi. Nel contempo lo zooplancton d'acqua fredda cede il posto a quello d'acqua calda: infatti alcune specie animali planctoniche di acqua calda del Mar dei Sargassi sfruttano le condizioni variabili degli anelli a nucleo freddo di età intermedia a tal punto da diventare più abbondanti qui che non altrove nel Mar dei Sargassi. La chiocciola marina d'acqua calda *L. inflata* ne è un esempio. Nell'aprile 1977 la sua quantità nel nucleo dell'anello Bob era ridotta, ma quattro mesi dopo era aumentata di un fattore 300. Nello stesso tempo la sua abbondanza nel Mar dei Sargassi era diminuita di un fattore 0,7, evidentemente a causa di un normale ciclo stagionale.

Altre specie di acqua calda si introducono più gradualmente in un anello a nucleo freddo in decadenza. A questo riguardo il genere planctonico *Stylocheiron* può essere del tutto tipico delle specie di acqua calda che occupano le nicchie ecologiche lasciate vacanti in un anello a nucleo freddo in decadenza dalle specie d'acqua di scarpata in diminuzione. Abbiamo spesso rinvenuto cinque specie di *Stylocheiron* nei campioni prelevati dall'Atlantico nordoccidentale. Diversamente da alcune altre specie planctoniche simili a gamberetti, esse mostrano una tendenza scarsa o nulla a spostamenti giornalieri verso l'alto o verso il basso. D'altra parte, esse mostrano un notevole grado di stratificazione verticale. Nel Mar dei Sargassi abbiamo trovato che il centro della distribuzione verticale di *S. elongatum* si trova tra 400 e 300 metri, quello di *S. affine* tra 250 e 200, quello di *S. abbreviatum* tra 200 e 100, quello di *S. carinatum* e di *S. suhmii* tra 100 metri e la superficie.

Nessuna di queste specie si trova in anelli molto giovani a nucleo freddo. L'anello Bob, per esempio, nell'aprile 1977 ne era completamente privo. Quattro mesi più tardi solo una di esse, *S. carinatum,* aveva raggiunto il centro, ma non le altre, neppure *S. suhmii* che vive quasi alla stessa profondità di *S. carinatum.* Dopo altri due o tre mesi anche le altre quattro specie avevano probabilmente raggiunto il centro. Gli anelli a nucleo freddo che abbiamo studiato, quando la loro età era compresa tra sei e nove mesi, contengono le cinque specie quasi alle stesse profondità che occupano nel Mar dei Sargassi. Evidentemente il trascorrere di sei mesi apporta una modificazione, nei 300 metri più superficiali di un anello a nucleo freddo, sufficiente per permettere a molte specie di acqua calda di svolgervi le loro normali funzioni vitali.

Gli anelli a nucleo caldo si formano secondo un processo che è in sostanza l'immagine speculare di quello che origina gli anelli a nucleo freddo. Di conseguenza ci si può aspettare che, in un anello a nucleo caldo in decadenza, avvengano trasformazioni biologiche contrarie a quelle che ho descritto per anelli a nucleo freddo come l'anello Bob? La risposta è incerta: la biologia degli anelli a nucleo caldo non è ancora stata studiata in dettaglio. Tuttavia, lo *status* fisiologico e biochimico di una specie animale planctonica dipende dalla quantità e dalla qualità del cibo disponibile nonché dalle condizioni del mare che influiscono sul tasso metabolico degli organismi. Probabilmente, le specie di acqua calda passate dal Mar dei Sargassi in un nucleo di un anello a nucleo caldo incontrano un'abbondanza crescente di cibo man mano che l'anello decade, ma devono far fronte a un tasso metabolico più basso dovuto alla diminuzione della temperatura dell'acqua.

Ho cercato di dimostrare in questo articolo che gli anelli meritano di essere studiati per loro stessi, ma anche e soprattutto perché offrono l'opportunità di studiare i meccanismi che regolano l'ecologia dell'oceano. Dopo tutto, la formazione, il decadimento e infine la scomparsa di un anello consentono di verificare quello che avviene quando una vasta popolazione di organismi marini che ha abbandonato il suo ambiente è soggetta a mutamenti idrografici e biologici al di là della sua normale esperienza. Gli stessi mutamenti hanno luogo in continuazione ai margini del normale habitat di una specie. Qui, però, è difficile determinare quali mutamenti particolari siano importanti. Negli anelli, sembra che i vari mutamenti (nella temperatura, nella salinità e così via) siano indipendenti, avvengano cioè a tassi diversi, così che un osservatore può solo sperare di apprendere quali siano quelli strettamente connessi all'alterazione dell'abbondanza di una specie. Questa occasione non ha precedenti nella biologia delle popolazioni. Uno studio sperimentale terrestre equivalente richiederebbe che un'area con un'estensione tripla di quella della Sardegna, con suoli, foreste, prati, fiumi, laghi e animali venisse controllata per un periodo di due anni e per di più in movimento.

Naturalmente ho parlato solo dell'Atlantico nordoccidentale. Si trovano anelli anche negli oceani attraversati da correnti come la Kuroshio, la corrente Agulhas, la corrente del Brasile, la corrente dell'Australia orientale e la corrente della Somalia. Altre forme di vortici si trovano presso la corrente della California e nei mari equatoriali. È del tutto verosimile che, ovunque scorrano forti correnti oceaniche, queste producano qualche forma di vortice. Ed è altrettanto verosimile che i vortici influenzino non solo l'idrografia del mare, ma anche la biologia. È probabile che i numerosi studi in corso portino presto a una miglior comprensione delle forze che governano gli oceani.

L'oncogène chiamato *src*, che induce un sarcoma nei polli, fa parte del piccolo genoma a RNA del «retrovirus» del sarcoma di Rous. Quando un retrovirus infetta una cellula, l'RNA virale viene copiato in un DNA a doppio filamento dall'enzima transcriptasi inversa, che è prodotto dal virus. Il DNA forma un anello e quindi si integra nel DNA della cellula ospite. Quando il DNA ospite viene trascritto in RNA dagli enzimi cellulari viene trascritto anche il DNA virale *(freccia tratteggiata)*. Una parte dell'RNA virale fornisce copie del genoma virale perché possano essere incluse nelle nuove particelle virali; una parte, invece, viene elaborata per produrre RNA messaggero, che viene tradotto in proteine virali. Parecchie di queste proteine sono poi incorporate nelle nuove particelle virali. Il prodotto del gene *src*, tuttavia, non è un componente della particella virale. È un enzima chiamato chinasi proteica, che si lega alla superficie interna della membrana plasmatica della cellula ed esegue la fosforilazione delle proteine cellulari, trasformando così la cellula normale in una cellula neoplastica.

I geni che provocano il cancro

Gli oncogèni sono stati trovati per la prima volta nei virus; ora si è visto che anche in normali cellule di vertebrati vi sono geni che, esprimendosi in modo anomalo, possono fare insorgere neoplasie

di J. Michael Bishop

È possibile conoscere la cellula neoplastica? Finora nessuno è riuscito a spiegare come una cellula normale controlli la propria crescita, può sembrare perciò azzardato pensare di poter decifrare le regole anomale che governano lo sviluppo di una cellula neoplastica. Eppure, la storia della biologia ha registrato molti casi in cui lo studio delle anomalie ha illuminato i processi vitali normali. I recenti sviluppi nel campo della ricerca oncologica hanno fornito un ulteriore sensazionale esempio: alcuni ricercatori hanno intuito, anche se vagamente, le linee essenziali degli eventi che possono indurre una crescita cancerosa. Sono stati identificati gli enzimi che li catalizzano e riconosciuti i geni che specificano la struttura di quegli enzimi.

Questi progressi derivano dagli studi compiuti sui virus che inducono tumori. Gli ultimi anni sono stati testimoni di appassionanti indagini sui virus cancerogeni negli esseri umani, indagini che purtroppo hanno avuto un successo limitato, tant'è che molti esperti ricercatori hanno incominciato a dubitare che mai si sarebbe riusciti a dimostrare che i virus sono una delle cause principali di cancro nell'uomo. Tuttavia, certi virus inducono tumori in altri animali e i ricercatori stanno studiandoli, nel tentativo di definire le fondamentali alterazioni della cellula, che sono responsabili dalla crescita neoplastica. È stato come scoprire l'oro.

Anche se i geni coinvolti nello sviluppo del cancro sono stati osservati per la prima volta in lavori sui virus, non sono originari dei virus. Anzi è emerso che non lo sono neppure nelle cellule normali e possono essere necessari per la vita di queste cellule così come lo sono per la crescita illimitata di quelle cancerose. Una via finale comune, dalla quale originano tutti i tumori, può far parte del corredo genetico di cui è dotata ogni cellula vivente.

I virus tumorali

Un virus è poco più di un pacchetto di informazioni genetiche, racchiuso in un involucro proteico. Le informazioni possono essere incorporate o nel DNA o nell'RNA (mentre nelle cellule degli organismi superiori l'archivio genetico è costituito sempre da DNA). Sia il DNA sia l'RNA sono lunghi filamenti costituiti da sequenze di quattro unità chimiche, chiamate nucleotidi, sequenze che costituiscono messaggi in codice divisi, per così dire, in unità distinte: i geni. Le istruzioni codificate nei geni sono trasmesse in vari modi. Il più delle volte la sequenza dei nucleotidi specifica l'ordine in cui gli amminoacidi vengono uniti per formare una particolare proteina, tipicamente un enzima o un elemento strutturale. I virus possono avere da meno di cinque fino a parecchie centinaia di geni mentre le cellule degli organismi più complessi hanno un genoma, o corredo genetico totale, di decine di migliaia di geni. La riproduzione dei virus imita i processi mediante i quali le cellule si accrescono e si dividono, ma la loro semplicità ne rende, in confronto alle cellule, molto più agevole lo studio e la composizione.

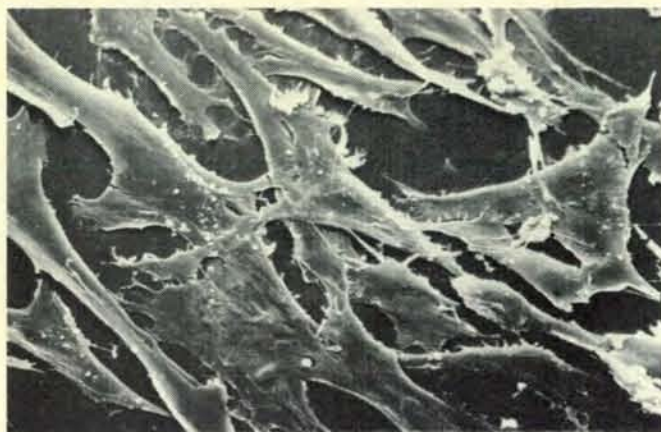
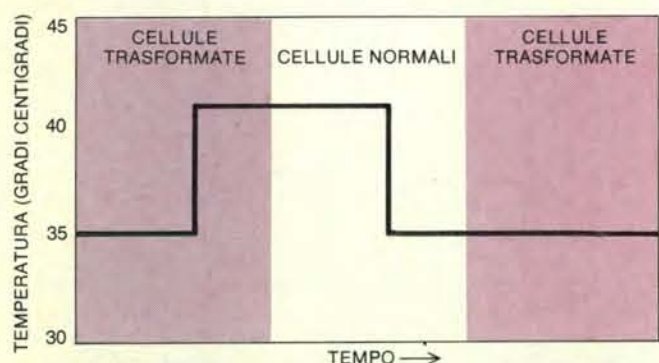
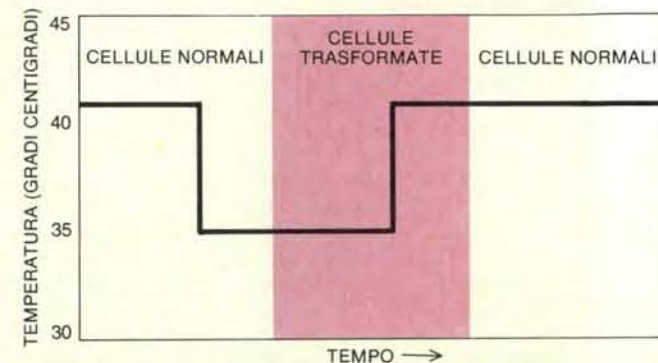
Nelle cellule, il DNA è trascritto in un filamento di RNA messaggero e questo RNA è tradotto in proteina. Un virus che infetta una cellula insinua la propria informazione genetica nella macchina cellulare, in modo che la cellula sintetizzi proteine virali specificate da geni virali. Le proteine sintetizzano molte copie del genoma virale, costruiscono nuove particelle virali ed eseguono qualunque altra informazione sia contenuta nei geni virali. In alcuni casi le istruzioni includono un comando che trasforma la cellula ospite in una cellula cancerosa.

L'esistenza di virus tumorali fu sospettata per la prima volta all'inizio del secolo. Una scoperta di importanza fondamentale si ebbe nel 1910, quando Peyton Rous del Rockefeller Institute for Medical Research dimostrò che un filtrato, privo di cellule, di quei tumori del pollo detti sarcomi poteva indurre sarcomi in altri polli. I suoi lavori non ottennero però una buona accoglienza e Rous finì con l'abbandonare le ricerche a causa della disapprovazione dei suoi colleghi. A distanza di decenni, l'esistenza del virus identificato per la prima volta da lui, così come di altri virus tumorali, fu inequivocabilmente provata mediante purificazione con tecniche fisiche e visualizzazione al microscopio elettronico. I virus tumorali divennero così elementi quotidiani nella ricerca oncologica. Nel 1966, all'età di 85 anni, Rous ricevette il premio Nobel.

Alcuni virus tumorali sono oncogeni (cioè inducono tumori) solo in animali che non sono i loro ospiti naturali, mentre altri virus tumorali lo sono proprio nei loro ospiti naturali. Tali differenze, comprese solo in parte, non rivestono tuttavia particolare interesse per il ricercatore. La possibilità di indurre tumori a volontà con un agente piuttosto semplice e ben definito ha costituito un grosso vantaggio per la ricerca sul cancro, anche se talvolta è necessario ricorrere a una combinazione innaturale tra virus e ospite.

La trasformazione

Molti virus tumorali hanno una proprietà particolarmente utile: in un terreno di coltura artificiale provocano modificazioni neoplastiche nelle cellule. Questa «trasformazione» delle cellule in coltura permette di esaminare l'interazione di un virus tumorale con una cellula ospite in condizioni controllate, evitando tutte le difficoltà che accompagnano gli esperimenti sugli animali. Tuttavia, è importante ricordare che vi sono alcuni virus tumorali che, pur essendo potenti agenti oncogeni negli animali, non trasformano le cellule in coltura.

La capacità o meno di un virus di apportare trasformazioni a cellule in coltura è connessa con il suo meccanismo di oncogenesi. Sono stati finora riconosciuti due meccanismi. Alcuni virus possiedono un solo gene, o talvolta pochi geni responsabili della loro capacità di indurre tumori che vengono chiamati «oncogèni». L'azione degli oncogèni virali è rapida e prevale sulla attività di tutti gli altri geni della cellula. La maggior parte dei virus con questo tipo di geni (forse tutti) è in grado di trasformare le cellule in coltura. La capacità di trasformazione è una prova

Nelle microfotografie al microscopio elettronico a scansione qui riprodotte, realizzate da G. Steven Martin dell'Università della California a Berkeley, è evidente la trasformazione di cellule in coltura operata dal virus del sarcoma di Rous. Fibroblasti (cioè cellule del tessuto connettivo) normali aderiscono alla superficie della capsula di Petri nella quale sono coltivati e si presentano in una configurazione appiattita e molto distesa *(a sinistra)*. Se vengono infettati con il virus del sarcoma di Rous, assumono invece un aspetto rotondeggiante e si ammassano in mucchi *(a destra)*; presumibilmente questa configurazione è causata dalla fosforilazione delle proteine cellulari da parte dell'enzima codificato dal gene *src*.

La mutazione condizionale di un gene porta a una inattivazione reversibile del gene o dell'enzima da esso codificato. Nel caso di un mutante di *src* sensibile alla temperatura, le cellule infettate sono trasformate se coltivate a 35 gradi centigradi ma a 41 ritornano o rimangono allo stato normale. La scoperta di questo tipo di mutante sottintende che un gene virale deve essere responsabile della trasformazione, che l'azione del gene è probabilmente mediata da un prodotto proteico e che è necessaria l'azione prolungata del gene perché avvenga la trasformazione.

non definitiva dell'esistenza di un oncogène in un virus. Altri virus sono privi di oncogèni e inducono tumori con mezzi più raffinati. Essi agiscono lentamente negli animali, in molti casi impiegando dai sei ai dodici mesi per dare origine a un tumore, a differenza dei virus con oncogèni, che impiegano solo alcuni giorni o settimane. E, inoltre, non trasformano le cellule in coltura.

Ambedue le forme di oncogenesi sono caratterizzate dalla persistenza del genoma virale nella cellula ospite fino a quando quest'ultima sopravvive. Nella maggior parte dei casi, il DNA virale si integra, o si attacca chimicamente, al DNA della cellula ospite; pare tuttavia che il genoma di alcuni virus tumorali si mantenga all'interno della cellula come unità separata, riproducendosi in modo indipendente. Per il momento sembra che la conservazione del genoma virale sia necessaria all'oncogenesi virale sia per mantenere l'influenza di un oncogène sulla cellula sia per sostenere gli effetti meno diretti dei virus che non portano un oncogène. I misteri dell'oncogenesi virale hanno talvolta suggerito l'ipotesi di un meccanismo tipo «batti e corri», in cui una infezione virale provvisoria innesca una sequenza di eventi che sfociano in un tumore, senza che nelle cellule tumorali persista alcuna traccia del virus. Si hanno oggi prove, sia pur minime, a favore di questi modelli.

I retrovirus

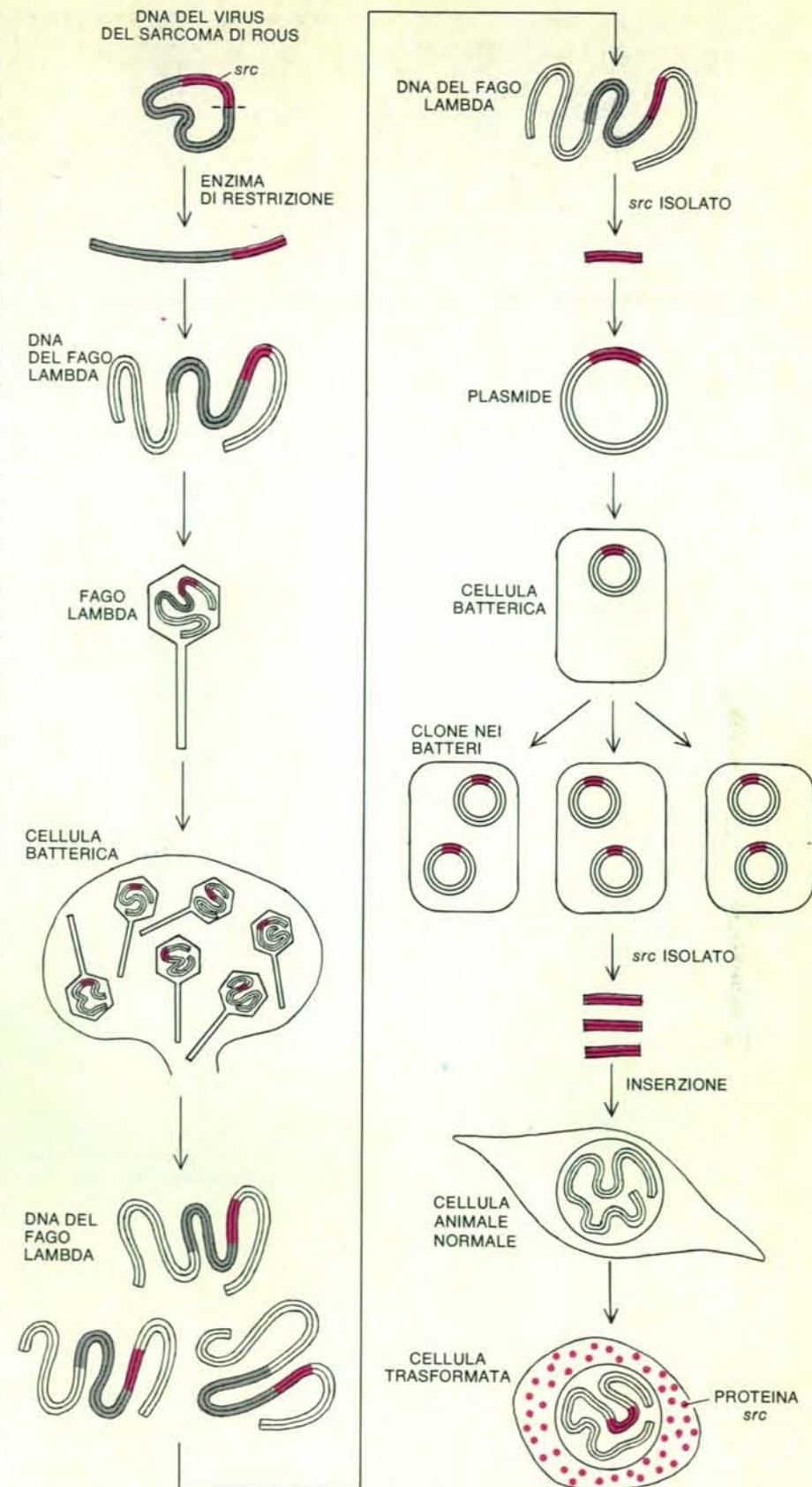
Il virus del sarcoma, scoperto da Rous, appartiene a una famiglia di virus noti come retrovirus, gli unici virus tumorali con genoma a RNA. Si tratta dei virus che hanno fornito, dell'oncogenesi, la spiegazione più logica di cui si possa disporre oggi. Tre caratteristiche dei retrovirus chiariscono la loro utilità nell'analisi dello sviluppo tumorale. Innanzitutto, sono stati trovati in un gran numero di specie di vertebrati e inducono molti tipi di tumori, che possono costituire modelli sperimentali per la maggior parte delle principali forme di cancro umano. In secondo luogo, è relativamente facile identificare e isolare gli oncogèni dei retrovirus e rintracciarne i prodotti; e ciò ha consentito di farsi una prima idea dei processi chimici che sono responsabili della crescita neoplastica. In terzo luogo, gli oncogèni dei retrovirus non sembrano essere componenti originarie del genoma virale; al contrario sembra che siano stati copiati da geni del vertebrato ospite, in cui il virus si è duplicato. C'è motivo di sospettare che i geni cellulari, da cui gli oncogèni dei retrovirus sono probabilmente derivati, siano essi stessi coinvolti nella produzione di tumori indotti da agenti diversi dai virus. Così, gli esperti di virologia dei tumori, impegnati nell'arcano tentativo di risalire all'origine evolutiva degli oncogèni, sono arrivati a conoscere meccanismi genetici che possono essere all'origine di molte forme di cancro.

I retrovirus devono questo nome a un aspetto del loro ciclo vitale che li rende unici in campo biologico: perché possano riprodursi, il loro RNA deve essere trascritto «a ritroso» nel DNA. Questo insolito processo viene realizzato da un enzima, chiamato transcriptasi inversa. La transcriptasi inversa fu scoperta per la prima volta nel 1970 nelle particelle virali del tipo del virus del sarcoma di Rous, da David Baltimore del Massachusetts Institute of Technology e da Satoshi Mizutani e Howard M. Temin dell'Università del Wisconsin. La scoperta fu importante sotto molti aspetti. Ha smentito l'erronea concezione, da molti sostenuta, che l'informazione genetica possa fluire solo dal DNA all'RNA. Ha innescato un'ondata travolgente di ricerche sui retrovirus, chiarendo il meccanismo fino ad allora oscuro della loro replicazione. Infine, ha fornito un reagente essenziale per la tecnologia dell'ingegneria genetica con DNA ricombinante, che si stava sviluppando allora.

Il ciclo vitale di un retrovirus è una meraviglia di cooperazione tra parassita e ospite. Il successo dell'infezione virale dipende dalla generosa ospitalità offerta dalla cellula, anche se il virus conserva un considerevole peso nel controllo degli eventi. Durante le prime ore dell'infezione, il genoma dell'RNA virale viene trascritto in DNA dalla transcriptasi inversa. Il DNA virale viene quindi integrato nel genoma della cellula, con il risultato che i geni virali si replicano assieme ai geni cellulari e vengono espressi dalla macchina cellulare.

In molti casi, un'infezione da retrovirus è innocua per la cellula. Il virus acquisisce una dimora nuova e potenzialmente stabile; vengono prodotte nuove particelle virali che abbandonano la cellula, eppure questa non subisce alcun danno. La collaborazione può, tuttavia, finire male in seguito all'uno o all'altro tipo di oncogenesi virale citati prima. Se il virus contiene un oncogène, l'attività di questo può portare la cellula a una crescita neoplastica. Se, invece, il virus è privo di oncogèni, l'integrazione del DNA virale può interferire con un gene cellulare in corrispondenza o in prossimità del punto di inserzione; in altre parole, l'inserzione può provocare una mutazione nel genoma della cellula ospite; e le mutazioni in certi siti possono generare neoplasie. L'induzione di tumori da parte degli oncogèni e quella provocata dalle conseguenze dell'integrazione sembrano, a prima vista, eventi del tutto diversi, ma mostrerò più avanti che sono, in realtà, intimamente legati.

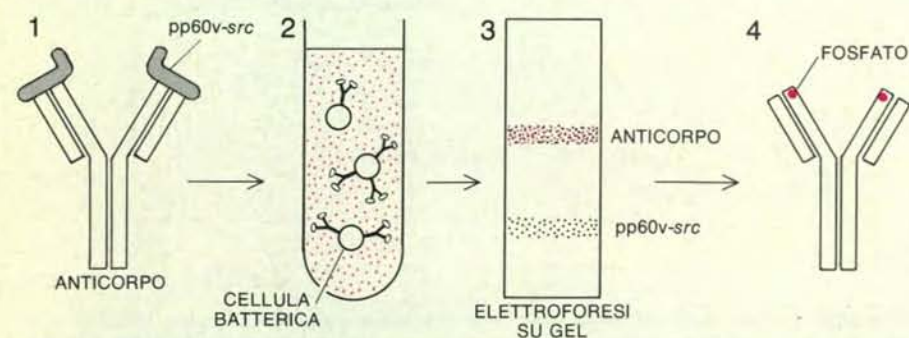
Il gene src

L'oncogène del virus del sarcoma di Rous fu sottoposto per primo all'analisi sperimentale. Una prima tappa importante è stata compiuta nel 1970, quando G. Steven Martin dell'Università della California a Berkeley ha identificato alcune mutazioni «condizionali», sensibili alla temperatura, che influiscono sulla capacità del virus di trasformare le cellule in coltura. Una mutazione condizionale è un potente strumento perché permette di inattivare reversibilmente un gene. Quando cellule in coltura, infettate con virus del sarcoma di Rous, sensibili alla temperatura, vengono mantenute in condizioni di temperatura «permissive», si trasformano. Quando la temperatura viene spostata verso valori più alti, «restrittivi», nell'arco di alcune ore le cellule riacquistano il loro aspetto normale, per tornare a trasformarsi di nuovo non appena la temperatura si abbassa. L'interpretazione del fenomeno è che, a una

L'oncogène virale viene purificato e viene saggiata la sua capacità di trasformare le cellule, con i metodi dell'ingegneria genetica. Il DNA circolare del virus del sarcoma di Rous è isolato da cellule infettate di recente, scisso mediante un enzima di restrizione e inserito nel DNA del fago lambda, un virus batterico. La crescita del fago nei batteri produce grossi quantitativi di DNA virale, che viene scisso da un enzima di restrizione: si ottiene così un frammento che contiene solo *src* e un po' di DNA ai lati. Il frammento viene inserito in un plasmide, che a sua volta viene introdotto in batteri per una ulteriore amplificazione. Così si possono excidere dai plasmidi, purificare e introdurre nelle cellule animali in coltura molte copie del frammento *src*; questo dirige la sintesi di una proteina virale che induce la trasformazione. L'esperimento è stato eseguito da William J. DeLorbe e da Paul A. Luciw per dimostrare che l'*src*, agendo da solo, dà origine a una neoplasia.

temperatura restrittiva, un gene mutato viene inattivato. La trasformazione, quindi, è dovuta all'azione di un gene, che, per poter mantenere lo stato canceroso, deve essere espresso con continuità. (Nella maggior parte dei casi, la temperatura elevata non agisce probabilmente sul gene in modo diretto. La mutazione altera invece la struttura del prodotto proteico del gene, con il risultato che l'attività della proteina viene danneggiata dalla temperatura restrittiva.)

Il gene osservato per la prima volta da Martin viene oggi chiamato *src* (da sarcoma, il tumore da esso indotto) ed è l'oncogène del virus del sarcoma di Rous. Poco tempo dopo esso fu meglio definito da Peter H. Duesberg di Berkeley e da Charles Weissmann, Martin Billeter e John M. Coffin dell'Università di Zurigo, che lavoravano tutti con ceppi di tale virus, isolati da Peter K. Vogt dell'Università della California meridionale. Si tratta di ceppi mutanti per «delezione», che hanno perso l'oncogène e quindi sono incapaci di indurre tumori o di trasformare le cellule in coltura. Duesberg, Weissmann e i suoi collaboratori hanno frammentato con l'enzima ribonucleasi i genomi di tali mutanti e di virus di tipo selvatico (con oncogèni). Determinando poi il frammento mancante nei mutanti, furono in grado di identificare l'oncogène come un segmento di RNA prossimo a una estremità del genoma del virus del sarcoma di Rous.

Negli ultimissimi anni, si è fatto ricorso alle nuove e potenti tecniche della ingegneria genetica per definire con maggior precisione gli oncogèni e saggiarne il potere cancerogeno. Con l'aiuto di una serie di enzimi, chiamati endonucleasi di restrizione, il DNA può oggi essere tagliato in frammenti in corrispondenza di siti specifici. Particolari frammenti possono essere fatti crescere in gran numero nei batteri, quindi isolati di nuovo e inseriti in cellule in coltura, dove i geni portati da quei frammenti di DNA possono essere espressi. In questo modo è possibile tagliare il DNA virale in segmenti che portano ciascuno un singolo gene per scoprire quali di questi segmenti provocano trasformazione. L'analisi del DNA del virus del sarcoma di Rous ha rivelato un unico gene capace di trasformare le cellule; tale gene codifica per un unico prodotto proteico. Ciò comporta che un unico gene, dirigendo la sintesi di un'unica proteina, può dar luogo ai cambiamenti che caratterizzano una cellula neoplastica. Conoscere questa proteina e il modo in cui agisce significa avere un'idea degli eventi che possono generare un tumore maligno.

Il prodotto dell'oncogène, una proteina designata come pp60v-*src*, è una catena di circa 250 subunità amminoacidiche. Si tratta di una chinasi proteica, un enzima che aggiunge gruppi fosfato alle proteine. Essa si lega alla membrana plasmatica delle cellule mediante un settore che è vicino all'estremità amminica (NH_2); il sito di fosforilazione si trova nell'altra metà della molecola.

L'attività di chinasi proteica della pp60v-*src* è stata scoperta per caso. Conigli nei quali sono stati indotti tumori da parte del virus del sarcoma di Rous producono anticorpi contro la proteina pp60v-*src* e questi anticorpi formano complessi con la proteina *(1)*. Tali complessi sono stati immobilizzati sulla superficie dei batteri ed è stata aggiunta una fonte di fosfato radioattivo *(puntini in colore) (2)*. Quando la pp60v-*src* e l'anticorpo sono stati separati mediante elettroforesi su gel, si è potuto accertare che la banda corrispondente all'anticorpo era radioattiva *(3)*; la pp60v-*src* immobilizzata aveva catalizzato il trasferimento del gruppo fosfato all'amminoacido tirosina in corrispondenza di un unico sito in ciascuna delle due catene pesanti dell'anticorpo *(4)*.

La proteina codificata dal gene *src* è nota in gran parte grazie al lavoro di Raymond L. Erikson e dei suoi collaboratori della School of Medicine dell'Università del Colorado. Essi cominciarono con l'identificare una proteina che viene sintetizzata in provetta in base alle istruzioni fornite dal genoma del virus del sarcoma di Rous di tipo selvatico, ma non in base a quelle fornite dal genoma di un mutante per delezione, privo del gene *src*. Poi, con il virus del sarcoma di Rous, indussero tumori in conigli e ottennero un discreto quantitativo di anticorpi di coniglio contro una presunta proteina *src*. Questi anticorpi si combinavano in maniera specifica con la proteina sintetizzata in provetta e, inoltre, con un'identica proteina presente nelle cellule trasformate dal gene *src*. Questi risultati hanno permesso di identificare con un certo grado di sicurezza una proteina che viene codificata dal gene *src* e che è responsabile dei suoi effetti. Essa è stata designata come pp60v-*src*: «pp» significa che è una fosfoproteina (una proteina alla quale sono attaccati gruppi fosfato), «60» si riferisce al suo peso molecolare di 60 000 dalton e «v-*src*» indica che la sua origine genetica è il gene virale *src*.

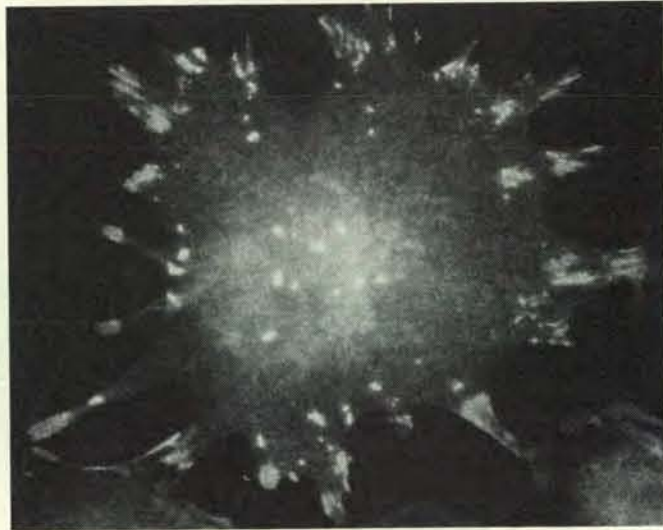
Un enzima del cancro

In che modo il prodotto proteico del gene *src* trasforma una cellula normale in una neoplastica? Quando tale proteina venne isolata, la domanda sembrò scoraggiante. Ma una prima risposta si ebbe rapidamente quando fu scoperto che la pp60-*src* è una chinasi proteica, un enzima che attacca gli ioni fosfato agli amminoacidi delle proteine nella reazione nota come fosforilazione. La scoperta fu compiuta da Erikson e dal suo collaboratore Mark S. Collett e, in modo indipendente, da Arthur Levinson, che lavora con Harold E. Varmus e con me nel laboratorio della School of Medicine dell'Università della California a San Francisco.

Poco tempo dopo Tony Hunter e Bartholomew M. Sefton del Salk Institute for Biological Studies hanno reso noto che la proteina pp60v-*src* attacca in modo specifico gli ioni fosfato all'amminoacido tirosina. Ciò poneva tale proteina al di fuori delle classi conosciute di chinasi proteiche, che fosforilano gli amminoacidi serina e treonina. La fosforilazione della tirosina si è rivelata una caratteristica comune degli enzimi codificati dall'oncogène, e, fatto ancor più sorprendente, svolge anche un ruolo nella regolazione della crescita delle cellule normali.

Non molti anni fa il fosfato era considerato dalla maggior parte dei biologi un

Questa microfotografia elettronica, scattata da M. C. Willingham e I. H. Pastan del National Cancer Institute, mostra dove è localizzata la maggior parte della proteina *src*. Cellule infettate con il virus del sarcoma di Rous sono state sezionate e trattate con un anticorpo di coniglio contro la pp60v-*src*. Una sostanza opaca al fascio di elettroni, la ferritina, era stata legata per via indiretta (cioè attraverso un secondo anticorpo) all'anticorpo: i punti scuri della ferritina evidenziano la localizzazione della pp60v-*src* sulle membrane plasmatiche tra due cellule confinanti.

Le placche adesive, che ancorano le cellule alle superfici, subiscono un'influenza diretta da parte della pp60v-*src*, come è dimostrato in queste microfotografie effettuate da Larry R. Rohrschneider del Fred Hutchinson Cancer Research Center di Seattle. Nella microfotografia al microscopio a interferenza e riflessione *(a sinistra)* le placche adesive, in contatto con la superficie, appaiono scure. Quando la stessa cellula viene trattata con un anticorpo contro la pp60v-*src*, che è marcata per via indiretta con un colorante fluorescente, una microfotografia al microscopio all'ultravioletto *(a destra)* mostra come la maggior parte della proteina sia localizzata proprio nelle placche adesive.

prodotto di scarso rilievo e il suo trasferimento alle proteine un evento modesto. Oggi, invece, risulta evidente che la fosforilazione delle proteine costituisce uno dei mezzi principali mediante i quali vengono controllate le attività dello sviluppo cellulare. Un enzima, fosforilando un certo numero di proteine, può provocare varie alterazioni nel funzionamento di una cellula. Nel caso della proteina pp60v-*src*, sono state proposte due modalità d'azione. L'enzima potrebbe fosforilare un'unica proteina, provocando un'improvvisa catena di eventi che concorrono a generare le proprietà di una cellula cancerosa; oppure, l'enzima potrebbe fosforilare numerose proteine, influendo direttamente sulle funzioni di ciascuna di loro e forse suscitando, successivamente, eventi secondari o, addirittura, una catena di eventi. Le poche informazioni di cui si dispone per il momento portano a ritenere la seconda ipotesi più probabile.

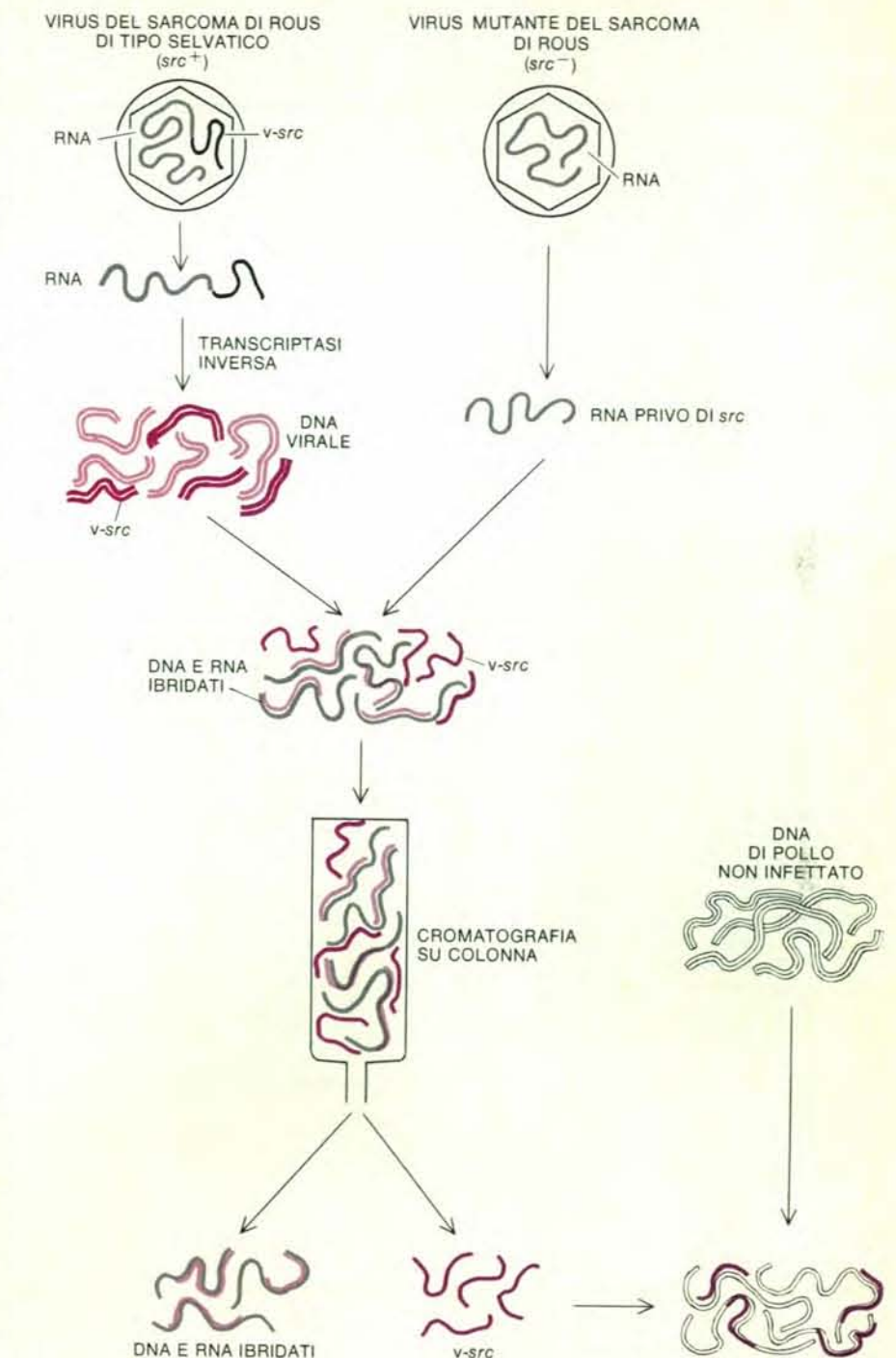
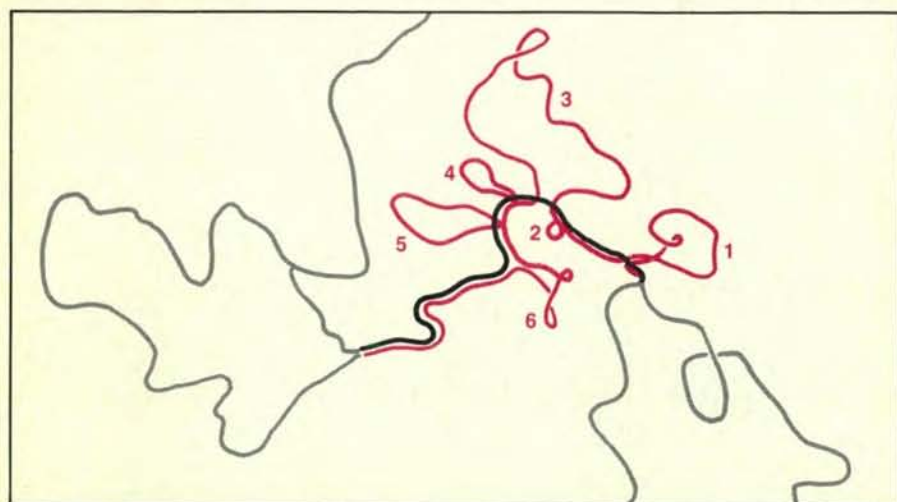
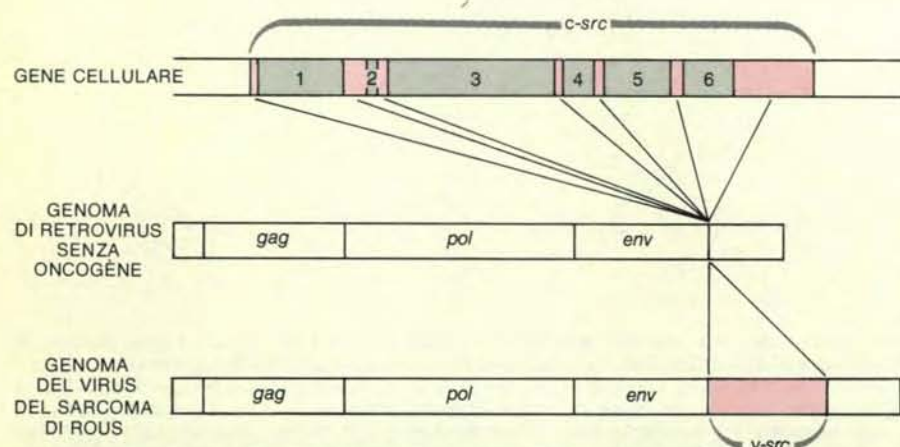
La fosforilazione delle subunità di tirosina nelle proteine cellulari può spiegare la capacità del gene *src* di indurre tumori? Hunter e collaboratori hanno mostrato che in una cellula trasformata ad opera del gene *src* la quantità di tirosina fosforilata aumenta all'incirca di dieci volte. L'aumento è considerato come una manifestazione dell'attività della proteina pp60v-*src*. Gli interrogativi cruciali a questo punto sono: quali proteine cellulari vengono fosforilate dall'enzima e quali sono le loro funzioni? Disponiamo solo di pochi indizi in proposito, e nessuno di questi può finora spiegare la crescita tumorale incontrollata, indotta dal gene *src*. La ricerca di bersagli per la proteina pp60v-*src* è in corso in molti laboratori.

Sito d'azione

Un modo di affrontare il problema consiste nel trovare dove, nella cellula, la proteina pp60v-*src* agisce, nella speranza di scoprire su quali proteine influisce e qual è la funzione di tali proteine. I primi studi inducevano a ritenere che i prodotti degli oncogèni virali potessero insediarsi nel nucleo cellulare, dove si sarebbero mescolati direttamente con l'apparato responsabile della replicazione del DNA, conducendo così la cellula ad una crescita incontrollata. Esperimenti effettuati da Hartmut Beug e da Thomas Graf del Max Planck Institut für Virusforschung di Tübingen hanno però dimostrato che gli effetti della proteina codificata dal gene *src* possono essere individuati anche in cellule da cui il nucleo è stato rimosso. Nessuna sorpresa si ebbe, pertanto, quando parecchi ricercatori trovarono che, nelle cellule trasformate, una piccola quantità (seppure ve n'era) della proteina pp60v-*src* si trovava nel nucleo. La maggior parte si trova, invece, all'altro estremo della cellula, legata alla membrana plasmatica, la sottile pellicola che avvolge la cellula e media le sue interazioni con il mondo esterno. Molti esperti di biologia cellulare hanno sostenuto che il controllo della crescita cellulare può avere origine in corrispondenza della membrana cellulare e delle strutture ad essa associate.

L'attento esame della membrana plasmatica di cellule trasformate dal gene *src* ha fornito la prima correlazione tra l'azione della pp60v-*src* su una specifica proteina cellulare e una delle tipiche modificazioni strutturali e funzionali osservate nelle cellule cancerose. Mediante tecniche specializzate che consentono di fotografare direttamente al microscopio, Larry R. Rohrschneider del Fred Hutchinson Cancer Research Center di Seattle è riuscito a dimostrare che la pp60v-*src* si concentra in placche adesive, zone della membrana che aderiscono a superfici solide. Nelle cellule neoplastiche esse vengono demolite e la minore adesività cellulare che ne risulta può facilitare il distacco dal tessuto di origine della maggior parte delle cellule di quel tipo e la creazione di metastasi in altre sedi.

I risultati di Rohrschneider portano a ritenere che la pp60v-*src* potrebbe demolire le placche adesive fosforilandone una, o forse più componenti proteiche. Seguendo quest'ipotesi, Sefton e S. J. Singer dell'Università della California a San Diego hanno mostrato che essa fosforila una unità tirosinica della vinculina, una proteina che è un costituente delle placche adesive normali e che, a seguito della trasformazione provocata dal gene *src*, si disperde nella cellula. Ragionevolmente la fosforilazione della vinculina accelera la demolizione delle placche adesive, ma il peso di tali eventi nel comportamento incontrollato delle cellule cancerose dev'essere ancora stabilito.

In passato si pensava che gli effetti cancerogeni dei virus potessero essere manifestazioni secondarie di geni virali, aventi come funzione principale quella di contribuire alla produzione di nuove particelle virali. Oggi è evidente che la replicazione dei retrovirus procede normalmente anche in assenza di oncogèni. Come si può spiegare allora la presenza su vasta scala degli oncogèni nei retrovirus e la loro evidente conservazione nel corso dell'evoluzione? Ricerche prolungatesi per un decennio hanno fornito una sorprendente risposta. Gli oncogèni dei retrovirus sono soltanto dei geni cellulari in un'altra guisa, passeggeri che gli animali in cui i virus si possono replicare hanno acquisito. La scoperta che anche le cellule hanno oncogèni ha implicazioni che vanno ben al di là della virologia tumorale.

Una sonda radioattiva utilizzata per ricercare il gene *src* in cellule normali è stata allestita da Dominique Stehelin. Dal virus del sarcoma di Rous è stato isolato dell'RNA contenente il gene v-*src* *(in alto a sinistra)*, che è stato copiato in DNA dalla transcriptasi inversa in presenza di precursori del DNA, marcati con elementi radioattivi. I frammenti di DNA radioattivi risultanti *(in colore)* sono stati denaturati (cioè sono stati separati in filamenti complementari) e mescolati con RNA estratto *(in alto a destra)* da un virus mutante privo di v-*src*. I singoli filamenti di DNA si «ibrideranno» con filamenti molto affini di DNA o di RNA. La maggior parte del DNA radioattivo si è così ibridato con l'RNA di un mutante per delezione, mentre i frammenti *src*, non potendo trovare RNA complementare, non hanno potuto farlo. Gli ibridi di DNA-RNA sono stati, quindi, separati mediante cromatografia su colonna dal DNA non ibridato dal *src*; si è ottenuta così una sonda *src* radioattiva. Quando quest'ultima è stata mescolata *(in basso a destra)* con DNA normale di pollo, si è ibridata, rivelando la presenza di un protooncogène cellulare, chiamato c-*src*.

Gli oncogèni cellulari e virali sono ben visibili in alto in una microfotografia elettronica, ottenuta da Richard C. Parker. Il DNA virale, che porta il gene *src*, e il DNA di pollo che porta, di tale gene, la versione cellulare, sono stati isolati e i DNA a doppio filamento sono stati denaturati. Quindi si sono fatti ibridare i singoli filamenti con altri filamenti affini. In questo caso, come appare nel disegno, il gene virale (*in nero*) e il gene cellulare (*in colore*) hanno formato un heteroduplex. (In grigio viene rappresentato il DNA estraneo necessario per clonare i geni.) Le anse nel filamento cellulare sono sei introni, cioè sequenze intercalari che interrompono le sequenze che codificano per le proteine (esoni) in molti geni cellulari, ma non nei geni dei retrovirus. Microfotografie di questo tipo hanno permesso di stabilire che gli oncogèni nelle cellule sono «nativi».

Il gene cellulare discontinuo c-*src* *(in alto)* è formato da esoni *(in colore)* e introni *(in grigio)*. Il gene cellulare era stato «scovato» in un retrovirus preesistente. Gli introni sono stati eliminati e gli esoni, saldati insieme, sono stati inseriti nel genoma virale *(al centro)* per completare il genoma del virus del sarcoma di Rous *(in basso)*. Oltre al gene *src* vi sono i geni *gag*, *pol* ed *env*, che codificano rispettivamente per la proteina del capside virale, per l'enzima transcriptasi inversa e per le punte glicoproteiche dell'involucro virale. Altri oncogèni dei retrovirus sembrano avere analoga origine.

L'origine degli oncogèni

Nel 1972 Dominique Stehelin, Varmus e io decidemmo di prendere in esame l'ipotesi riguardante l'oncogène proposta da Robert J. Huebner e da George J. Todaro del National Cancer Institute. Nel cercare un unico meccanismo che spiegasse l'induzione del cancro a opera di differenti agenti, Huebner e Todaro avevano ipotizzato che gli oncogèni dei retrovirus fossero una parte del bagaglio genetico di tutte le cellule, forse acquisita attraverso un'infezione virale in una fase iniziale dell'evoluzione. Finché rimangono quiescenti, gli oncogèni sarebbero innocui. Tuttavia, se attivati da un agente cancerogeno, potrebbero trasformare le cellule normali in cellule cancerose. Concludemmo che, se l'ipotesi era giusta, avremmo potuto trovare il gene *src* nel DNA di cellule normali.

Il DNA dei vertebrati contiene decine di migliaia di geni. Per individuare, fra gli innumerevoli geni, il gene *src*, Stehelin ideò un efficace strumento: del DNA radioattivo copiato a opera della transcriptasi inversa esclusivamente a partire dal suddetto gene *(si veda l'illustrazione nella pagina precedente)*. Il DNA così ottenuto veniva poi utilizzato per cercare il DNA cellulare che avesse una sequenza nucleotidica analoga a quella del gene *src*. La ricerca venne realizzata mediante ibridazione molecolare in cui catene di un acido nucleico (DNA o RNA) si ibridano, oppure formano complessi con acidi nucleici affini. Fummo lieti di scoprire, con non poco stupore, che la copia del gene *src* di Stehelin poteva ibridarsi con DNA di polli e altri uccelli non infetti. Deborah H. Spector, unitasi al nostro gruppo, riuscì a trovare del DNA affine al gene *src* anche nei mammiferi, uomo compreso, e nei pesci. Concludemmo allora che tutti i vertebrati possiedono probabilmente un gene affine al *src*, e ciò sembrava confermare l'ipotesi riguardante l'oncogène di Huebner e Todaro.

A un'indagine più minuziosa, però, il gene che avevamo scoperto nei vertebrati non risultò affatto essere un gene di retrovirus. Si trattava, invece, di un gene cellulare, oggi chiamato c-*src*. La prova determinante che ci ha portato a questa conclusione è venuta dal fatto che l'informazione del c-*src*, che codifica per una proteina, è suddivisa in parecchi domini, chiamati esoni, intervallati da regioni note come introni. Una configurazione discontinua di questo tipo è caratteristica dei geni della cellula animale, ma non dei geni del retrovirus. A parte gli introni, le versioni del c-*src* che si trovano nei pesci, negli uccelli e nei mammiferi sono tutte strettamente correlate al gene virale v-*src* e fra di loro. Se ne deduce che il gene *src* dei vertebrati si è conservato attraverso lunghi periodi evolutivi senza sostanziali modificazioni, e che quindi costituisce un vantaggio per la specie in cui è presente.

Il mistero che circonda il gene c-*src* si è fatto ancor più fitto con la scoperta che non solo esso è presente nelle cellule normali, ma è anche attivo, in quanto vie-

ne trascritto nell'RNA messaggero che viene a sua volta tradotto in proteina. L'ibridazione molecolare con la copia radioattiva del v-*src* di Stehelin evidenziò dapprima l'RNA, sia nelle cellule di uccello sia in quelle di mammifero. La proteina era, invece, più sfuggente, soprattutto perché, nella maggior parte delle cellule, viene sintetizzata in piccolissime quantità. Avemmo successo quando, come altri, ci mettemmo a cercare la proteina cellulare con anticorpi preparati originariamente per la ricerca della proteina trasformante virale, pp60v-*src*. La proteina cellulare isolata con l'aiuto di questi anticorpi si dimostrò praticamente indistinguibile dalla proteina virale, e venne quindi chiamata pp60c-*src*. Le due proteine hanno dimensione e struttura chimica simili; ambedue catalizzano la fosforilazione della tirosina e sono saldamente legate alla membrana plasmatica (di cellule modificate, nel caso della proteina pp60v-*src*, di cellule normali nel caso della proteina pp60c-*src*). È come se le due proteine fossero progettate per lo stesso scopo, anche se l'una è una proteina virale che provoca il cancro e l'altra è una proteina di cellule normali.

Gli oncogèni cellulari

I risultati relativi all'*src* costituiscono il primo abbozzo di una generalizzazione, la cui estensione e il cui significato devono ancor oggi essere stabiliti. Dei diciassette oncogèni di retrovirus finora identificati, ne sono noti sedici che hanno «parenti stretti» nei genomi normali delle specie di vertebrati. La maggior parte di questi oncogèni cellulari affini agli oncogèni virali obbediscono alle stesse regole dedotte per la prima volta per il gene *c-src*. Hanno l'organizzazione strutturale dei geni cellulari piuttosto che quella dei geni virali, sembrano aver superato lunghi periodi di evoluzione e sono attivi nelle cellule normali. Per fornire una spiegazione di queste caratteristiche e della notevole somiglianza tra gli oncogèni dei retrovirus e i loro «parenti» delle cellule normali, la maggior parte dei virologi ha avanzato l'ipotesi che gli oncogèni dei retrovirus siano copie dei geni cellulari. Sembra che essi si siano aggiunti a genomi preesistenti di retrovirus in un passato non troppo lontano. In che modo e perché i retrovirus abbiano copiato i geni cellulari non lo sappiamo, ma vi sono motivi per pensare che questo processo continui ancor oggi e che sia persino possibile ripercorrere le tappe di questo processo in laboratorio.

Secondo l'ipotesi del dosaggio, un oncogène cellulare dirige la sintesi di una certa quantità di un prodotto proteico normale, necessario per lo sviluppo normale *(fila in alto)*; la trasformazione in neoplasia deriva da una superproduzione di tale prodotto. Nel caso di un'infezione da parte di un retrovirus tumorale *(1)*, la superproduzione è diretta, sotto controllo virale, dall'oncogène virale. Anche l'oncogène cellulare induce, potenzialmente, i tumori, come è stato dimostrato da un esperimento con DNA ricombinante *(2)*. Se un gene regolatore virale viene attaccato all'oncogène cellulare e questa forma attivata del gene viene reinserita nelle cellule in coltura, possono essere sintetizzate quantità eccessive del prodotto dell'oncogène, con conseguente trasformazione della cellula. Un analogo processo può spiegare la crescita neoplastica indotta da un'ampia gamma di agenti diversi dai retrovirus *(3)*: se il DNA della cellula viene danneggiato da un mutageno o da un altro agente cancerogeno in un modo che faccia aumentare l'attività dell'oncogène cellulare, il suo prodotto può essere sintetizzato in quantità eccessive.

I geni dei vertebrati da cui gli oncogèni dei retrovirus sono, a quanto pare, derivati furono chiamati in un primo momento protooncogèni sia per sottolineare il loro significato evolutivo sia per rendere implicito che i geni cellulari stessi hanno un potenziale oncogeno. Oggi invece è chiaro che essi possiedono realmente questo potenziale, che è costituito dagli oncogèni cellulari. Le indagini che spiegano una simile attribuzione hanno preso l'avvio da questo interrogativo: se gli oncogèni dei retrovirus sono semplici copie di geni presenti nelle cellule normali, come si possono spiegare gli effetti devastanti dei geni virali sulle cellule infette? Sono state proposte due spiegazioni. L'ipotesi della mutazione propone che gli oncogèni virali differiscano dai loro progenitori cellulari in modo impercettibile, ma determinante, a causa di mutazioni introdotte nel momento in cui i geni cellulari sono stati copiati per dare origine al genoma del retrovirus. Per esempio, le attività enzimatiche chiaramente simili delle proteine pp60v-*src* e pp60c-*src* potrebbero in realtà avere differenti bersagli nella cellula e, di conseguenza, effetti molto diversi sul comportamento cellulare. L'ipotesi del dosaggio suggerisce invece che gli oncogèni dei retrovirus agiscano «brutalmente», sovraccaricando le cellule con un eccesso di proteine essenzialmente normali, che svolgono funzioni normali. Secondo questa prospettiva la genesi del cancro a opera degli oncogèni dei retrovirus dipende dalla quantità delle proteine virali più che da una qualunque loro proprietà particolare.

È troppo presto per stabilire quale di queste ipotesi sia corretta, ma i primi dati favoriscono quella del dosaggio. Innanzitutto, le dosi delle proteine trasformanti dei retrovirus sono indiscutibilmente elevate. I segnali che dirigono l'attività dei geni dei retrovirus sono piuttosto potenti e quindi la quantità di proteina prodotta da un oncogène virale è di gran lunga superiore a quella generalmente prodotta dal corrispondente gene cellulare. È evidente che la cellula può essere sopraffatta. Conferme più determinanti sono emerse nel tentativo di saggiare una delle previsioni fondamentali dell'ipotesi del dosaggio: se gli oncogèni dei retrovirus e quelli cellulari hanno effettivamente un'identica funzione dovrebbe essere possibile identificare delle circostanze in cui possano essere gli stessi geni cellulari a indurre la formazione di neoplasie.

L'oncogenesi da geni cellulari

La prima conferma di questa previsione si è avuta dagli importanti esperimenti condotti da Hidesaburo Hanafusa e collaboratori della Rockefeller University, i quali trovarono ceppi del virus del sarcoma di Rous che avevano perduto ampi tratti del gene *src* (ma non tutto il gene) e che pertanto erano incapaci di indurre negli animali da laboratorio il caratteristico sarcoma. Quando Hanafusa iniettò dei virus «mutilati» nei polli e, successivamente, recuperò le particelle virali sintetizzate nelle cellule infette, scoprì con stupore che il gene v-*src* del virus era stato ricostituito. Evidentemente, durante lo sviluppo del virus nei polli, il materiale genetico del gene c-*src* si era ricombinato con il genoma virale. Il virus portatore del gene ricostituito era di nuovo perfettamente in grado di provocare tumori, anche se aveva appena acquisito i tre quarti del suo oncogène da un gene cellulare. Hanafusa fu in grado di ripetere più volte questo straordinario esperimento, sia nei polli sia nelle quaglie. I suoi risultati rafforzarono l'idea che il c-*src* e il v-*src* svolgano le stesse funzioni, ma molti esperti di virologia tumorale, mancando prove più dirette di un potenziale cancerogeno dei geni cellulari, rimasero scettici.

Ora queste prove ci sono. I gruppi di ricerca di George F. Vande Woude e di Edward M. Scolnick del National Cancer Institute hanno sfruttato le tecniche di ingegneria genetica per isolare tre oncogèni cellulari (uno di topo e due di ratto) e dimostrare per via diretta che essi possono indurre una crescita neoplastica in cellule in coltura. Essi attaccarono ai geni cellulari un gene regolatore virale, un segnale codificato dal DNA, che contribuisce a regolare l'espressione di un gene vicino. In conformità con l'ipotesi del dosaggio, quando si introduceva nelle cellule il complesso *src*-gene regolatore, alcune di queste venivano trasformate, proprio come se avessero ricevuto un oncogène virale, mentre in realtà avevano ricevuto un gene cellulare posto agli ordini di un virus, per intensificarne l'attività. Inoltre, come si poteva dimostrare, le cellule trasformate dai due oncogèni cellulari di ratto producevano grandi quantità di proteine codificate da quei geni, sempre in accordo con l'ipotesi del dosaggio.

Perché la sovrabbondanza di una proteina normale dovrebbe provocare effetti tanto devastanti? Si potrà rispondere con sicurezza a questa domanda solo dopo aver compreso il ruolo degli oncogèni cellulari nel normale funzionamento delle cellule. Forse gli oncogèni cellulari fanno parte di una rete di controlli, caratterizzati da un delicato equilibrio, che regolano la crescita e lo sviluppo delle cellule normali. Un'attività eccessiva da parte di uno solo di questi geni potrebbe spostare l'equilibrio della regolazione verso una crescita continua.

Si hanno prove che le attività dirette dagli oncogèni virali e cellulari contribuiscono al controllo della crescita delle cellule normali. Inizialmente pareva che la fosforilazione della tirosina a opera della proteina pp60v-*src* fosse un processo anomalo e che la sua natura estranea potesse essere alla base della risposta cancerosa all'*src*, ma, quando Stanley Cohen della School of Medicine della Vanderbilt University scoprì che tale fosforilazione aveva un ruolo normale nella vita delle cellule, le cose si capovolsero. Avendo scoperto e purificato un piccolo

Il concetto di gene del cancro, che trova sostegno nei dati sugli oncogèni e su altre prove preliminari, suggerisce una spiegazione unificatrice per varie forme di cancerogenesi. L'elemento centrale comune è un gruppo di geni cellulari necessari per una crescita e uno sviluppo normali. Trapiantato in un genoma di retrovirus *(a sinistra)*, uno di questi geni diventa un oncogène. Può aversi cancro anche se il gene cellulare è colpito da un altro agente mutageno o cancerogeno *(a destra)*.

ONCOGÈNE	SPECIE D'ORIGINE	TIPO DI TUMORE	PROTOONCOGÈNE NEL DNA DEI VERTEBRATI	PRODOTTO DEGLI ONCOGÈNI		
				CHINASI PROTEICA	FOSFORILA LA TIROSINA	SI LEGA ALLA MEMBRANA
v-src	POLLO	SARCOMA	SÌ	SÌ	SÌ	SÌ
v-fps	POLLO	SARCOMA	SÌ	SÌ	SÌ	SÌ
v-yes	POLLO	SARCOMA	SÌ	SÌ	SÌ	?
v-ros	POLLO	SARCOMA	SÌ	SÌ	SÌ	?
v-myc	POLLO	CARCINOMA, SARCOMA, LEUCEMIA	SÌ	?	?	?
v-erb	POLLO	LEUCEMIA, SARCOMA	SÌ	?	?	?
v-myb	POLLO	LEUCEMIA	SÌ	?	?	?
v-rel	TACCHINO	LINFOMA	SÌ	?	?	?
v-mos	TOPO	SARCOMA	SÌ	?	?	?
v-bas	TOPO	SARCOMA	SÌ	?	?	?
v-abl	TOPO	LEUCEMIA	SÌ	SÌ	SÌ	SÌ
v-ras	RATTO	SARCOMA, LEUCEMIA	SÌ	SÌ	?	SÌ
v-fes	GATTO	SARCOMA	SÌ	SÌ	SÌ	?
v-fms	GATTO	SARCOMA	SÌ	SÌ	?	?
v-sis	SCIMMIA	SARCOMA	SÌ	?	?	?

Gli oncogèni dei retrovirus sono elencanti assieme alle loro specie d'origine e ai tumori che inducono e che costituiscono modelli sperimentali per la maggior parte delle forme principali di cancro umano. Ciascuno di questi 15 oncogèni virali è correlato, a quanto pare, con un protooncogène presente nel DNA dei vertebrati. Il prodotto di almeno otto dei quindici geni è una chinasi proteica, un enzima fosforilante. È stato dimostrato che sei delle otto chinasi fosforilano la tirosina; almeno quattro si legano alla membrana plasmatica della cellula neoplastica.

«fattore di crescita epidermico», il cui legame con la superficie cellulare stimola la sintesi del DNA e la divisione cellulare, egli si mise alla ricerca del meccanismo che permettesse la trasmissione del segnale dall'esterno all'interno della cellula e, come prima cosa, poté dimostrare che il legame del fattore di crescita epidermico con le cellule porta alla fosforilazione delle proteine. Incoraggiato dai risultati ottenuti con la proteina pp60v-*src*, arrivò a scoprire che la fosforilazione indotta dal fattore di crescita epidermico riguarda in modo specifico la tirosina. In seguito, altri ricercatori hanno dimostrato che alcune proteine vengono fosforilate, come risposta al fattore di crescita epidermico, dalla proteina pp60v-*src*. Uno stimolatore normale della divisione cellulare (il fattore di crescita epidermico) e uno stimolatore anormale (la proteina pp60v-*src*) sembrano così agire in sintonia. È implicito che la fosforilazione della tirosina a opera della proteina pp60v-*src* sia coinvolta nella regolazione della crescita delle cellule normali.

Alla ricerca di una teoria unificata

Pare che i retrovirus non siano una delle cause principali del cancro umano. Ciononondimeno sembrano aver indicato la strada verso i meccanismi essenziali all'origine della malattia. Si ritiene in generale che il cancro insorga come conseguenza di un danno a carico del DNA, ma la natura esatta di tale danno è ancora oggetto di discussione. Come potrebbe, per esempio, provocare la crescita neoplastica? I tentativi di trovare una risposta valida per tutte le forme di cancro hanno, in gran parte, chiamato in causa l'esistenza di «geni del cancro», ossia di componenti del genoma cellulare normale, la cui attività, consentita o addirittura esaltata da agenti cancerogeni di vario tipo, è responsabile del comportamento anomalo delle cellule neoplastiche. In questo schema, i geni del cancro sono visti non come elementi estranei, ma come geni normali, in realtà essenziali, piombati in un'attività frenetica. Il danno provocato da un agente cancerogeno trasforma un amico in nemico, forse agendo direttamente sul gene del cancro o forse mutilando un secondo gene, che normalmente è preposto a sorvegliarne l'attività.

I ricercatori di genetica medica hanno individuato gli effetti dei geni del cancro anni fa, quando identificarono per la prima volta famiglie i cui componenti ereditano una predisposizione a qualche particolare forma di cancro. Sembra oggi che gli esperti di virologia tumorale abbiano messo direttamente le mani su tali geni nella forma di oncogèni cellulari. Nella loro forma virale, questi geni sono cancerogeni e i risultati sia di Vande Woude sia di Scolnick fanno ritenere che anche gli oncogèni cellulari possano trasformare le cellule. È facile, pertanto, immaginare che i geni del cancro e gli oncogèni cellulari, evidenziati dai retrovirus, siano la stessa cosa. L'ipotesi dell'oncogène è stata riproposta sostituendo ai geni virali i geni cellulari. L'ipotesi del dosaggio serve a spiegare perché l'aumento di attività di un gene cellulare normale potrebbe provocare il cancro.

Ulteriori prove a sostegno di queste ipotesi si hanno dallo studio dei retrovirus di pollo che inducono linfomi, tumori letali del sistema immunitario. I virus del linfoma di pollo non hanno oncogèni. Perché allora provocano dei tumori? William S. Hayward e Benjamin G. Neel della Rockefeller University e Susan M. Astrin dell'Institute for Cancer Research di Fox Chase, in Pennsylvania, hanno scoperto che, nei tumori indotti dal virus del linfoma del pollo, il DNA virale è quasi sempre inserito nel DNA cellulare nell'immediata vicinanza di un oncogène cellulare (non il c-*src*, ma un oncogène scoperto di recente, noto come c-*myc*). Come evidente conseguenza dell'inserzione, l'espressione dell'oncogène cellulare viene notevolmente amplificata.

Questi risultati concordano perfettamente con il concetto di geni del cancro. L'inserimento del DNA del virus del linfoma nel genoma dell'ospite è un fenomeno analogo alla mutagenesi o ad altre forme di alterazione, provocate da agenti cancerogeni di vari tipi. L'inserimento stimola chiaramente l'attività di un gene, che è noto per essere oncogeno, quando è presente (come v-*myc*) in un diverso retrovirus di pollo. L'azione stimolata dell'oncogène cellulare sembra essere responsabile, almeno in parte, della genesi dei tumori. I retrovirus senza oncogèni inducono vari tipi di tumori; identificando il sito in cui, in alcuni di questi tumori, il genoma virale si inserisce nel DNA cellulare, i virologi possono arrivare a scoprire geni del cancro non ancora identificati con altri mezzi.

La scoperta dei geni del cancro (nella forma di oncogèni cellulari) mediante retrovirus è stata piacevole e inattesa. Ma, a quanto pare, i ricercatori non si ritengono ancora soddisfatti del ritmo con cui i retrovirus ci offrono nuovi oncogèni dall'interno della cellula. Robert A. Weinberg del Massachusetts Institute of Technology e Geoffrey Cooper della Harvard Medical School hanno esteso la ricerca di geni del cancro oltre il campo della virologia tumorale. Essi hanno dimostrato che i frammenti di DNA aventi la lunghezza di un gene, isolati da alcuni tumori (non indotti da virus), possono trasmettere la capacità di indurre una crescita neoplastica se introdotti in cellule in coltura, prima normali.

Weinberg e Cooper hanno chiaramente trovato un modo di trasferire geni del cancro attivi da una cellula a un'altra. Essi dispongono di prove che differenti geni del cancro sono attivi in diversi tipi di tumori; sembra pertanto che il loro modo di affrontare il problema possa ampliare considerevolmente il repertorio di geni del cancro sui quali studiare. Nessuno dei geni del cancro finora scoperti da Weinberg e Cooper risulta identico a uno qualsiasi degli oncogèni noti. Eppure è probabile che esista un'unica grande famiglia di oncogèni cellulari. Se è davvero così, lo studio dei retrovirus e i procedimenti messi a punto da Weinberg e Cooper, finalmente, dovrebbero cominciare a fornirci dei campioni comuni da quest'unica famiglia.

Una via finale comune

Le cellule normali possono portare i semi della loro distruzione sotto forma di geni del cancro, le cui attività possono rappresentare la via finale comune, mediante la quale agiscono molti agenti cancerogeni. I geni del cancro possono non essere ospiti indesiderati, bensì componenti essenziali dell'apparato genetico cellulare, che tradiscono la cellula solo quando la loro struttura o il loro controllo si altera per azione degli agenti cancerogeni. Alcuni di questi geni perlomeno sembrano aver fatto la loro comparsa nei retrovirus, dove possono essere facilmente identificati, manipolati e caratterizzati.

Quello che si è appreso dagli oncogèni rappresenta un primo lume dietro la cortina che per tanto tempo ha oscurato i meccanismi del cancro. In un certo senso, la prima immagine è deludente, perché i meccanismi chimici che sembrano deviare la cellula non sono di tipo diverso da quelli che agiscono in una cellula normale. Ciò fa pensare che le progettate strategie terapeutiche possano rimanere al punto morto in cui si trovano oggi. Non avrà alcuna utilità inventare metodi per impedire le attività responsabili della crescita neoplastica se queste stesse attività sono anche necessarie per la sopravvivenza delle cellule normali.

In qualunque modo si concluda la saga degli oncogèni, in ogni caso essa offre alcuni insegnamenti per chiunque sia interessato alle ricerche sul cancro. Lo studio di virus non immediatamente collegato all'ambito umano ha permesso di arrivare a strumenti particolarmente utili per lo studio delle malattie umane. La virologia tumorale è riuscita a sopravvivere, nonostante non abbia saputo trovare nel cancro umano i numerosi agenti virali di cui andava alla ricerca. La questione che si pone oggi non è se i virus possano provocare tumori nell'uomo (come in alcuni casi è probabile), ma piuttosto quanto si può apprendere dalla virologia tumorale sui meccanismi mediante i quali i tumori insorgono.

Il laser nei processi lavorativi

I dispositivi laser, già impiegati con successo in settori come foratura, taglio, saldatura e allegamento superficiale, possono portare soluzioni nuove anche ad altri aspetti inerenti al sistema prodotto-produzione

di Aldo V. La Rocca

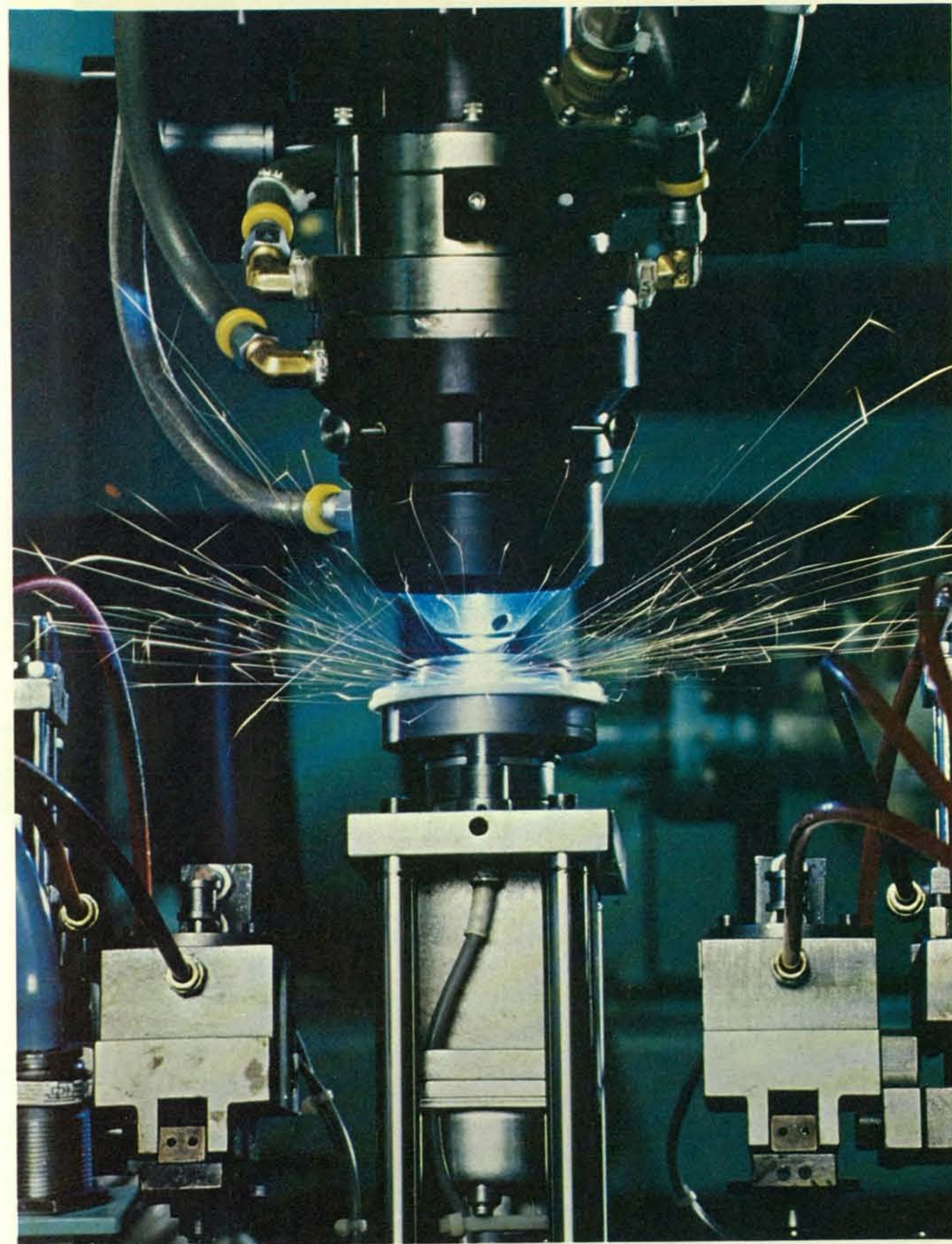
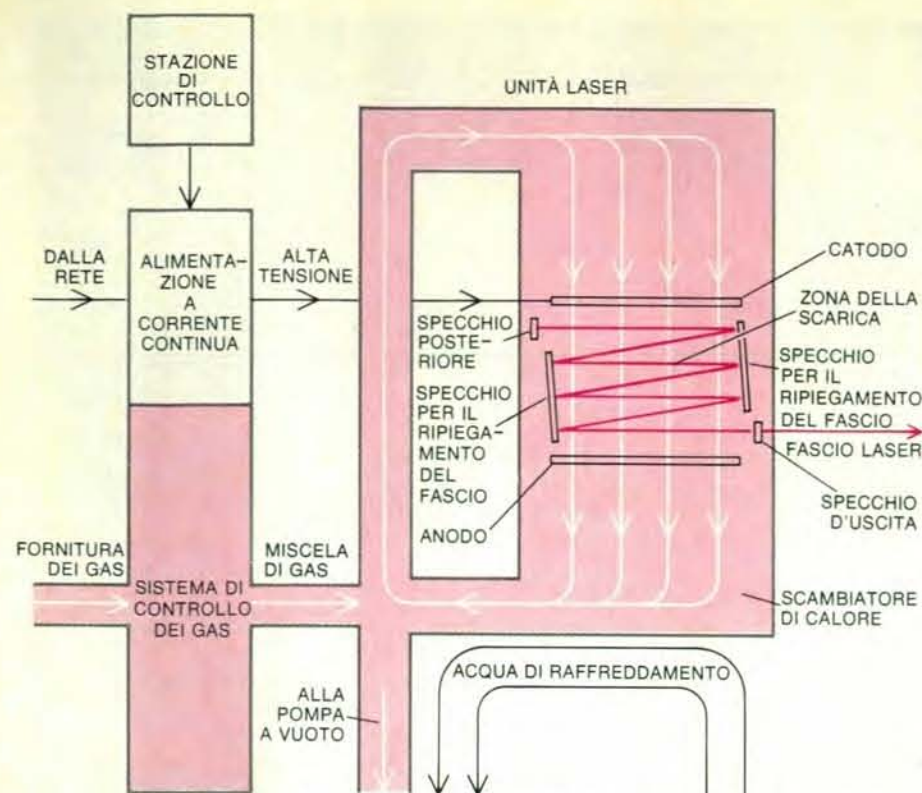
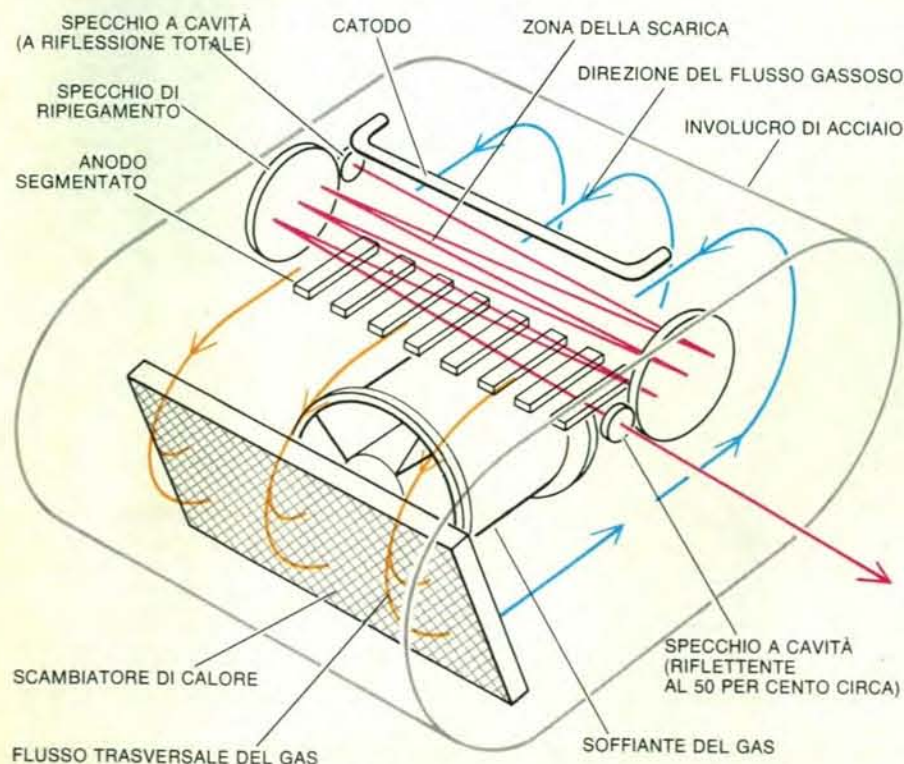
Una delle principali caratteristiche della radiazione laser è la sua intensità. «Laser» significa amplificazione della luce mediante emissione stimolata di radiazioni («light amplification by stimulated emission of radiation») e l'amplificazione genera una valanga di fotoni (quanti di radiazione elettromagnetica) entro un fascio ben collimato. Con l'invenzione del laser negli anni cinquanta ci si rese subito conto che i fasci laser di grande intensità avrebbero reso possibile il deposito di grandi quantità di potenza sui materiali nei processi lavorativi. Tale possibilità è ora diventata una tecnologia in pieno sviluppo. Durante lo scorso decennio i laser ad alta potenza sono stati impiegati in molti processi lavorativi: la saldatura di componenti automobilistici, di dispositivi elettronici e di strumenti medici; il trattamento termico di particolari automobilistici e aeronautici per migliorarne le caratteristiche superficiali; il taglio di lamiere nell'industria della formatura e tranciatura e la foratura di minuscoli orifizi di raffreddamento (da 0,007 a 0,05 pollici, ossia da 0,175 a 1,25 millimetri) sui particolari aeronautici. In tutte queste operazioni i sistemi laser hanno reso le linee di produzione maggiormente efficienti e hanno portato alla riduzione dei costi. Se si considera il gran numero di applicazioni attualmente in fase di realizzazione, sembra certo che nel corso del prossimo decennio la tecnologia laser si affermerà in molti altri processi produttivi.

Nei processi lavorativi il laser serve essenzialmente quale dispositivo capace di applicare un flusso estremamente elevato di energia sulla superficie del pezzo da lavorare. In tal senso esso comporta vantaggi significativi rispetto alle tradizionali sorgenti di calore quali la fiamma, la torcia, l'arco elettrico e il getto di plasma. Tra l'altro, il laser offre il vantaggio di un prodotto di alta qualità (in termini di migliori prestazioni e della riduzione del numero di particolari che devono essere rilavorati o ridotti in rottami); la riduzione degli investimenti per i materiali, la mano d'opera e la lavorazione; un'alta produttività (con conseguenti riduzioni di spazio e costi di ammortamento); un migliore ambiente di lavoro, la flessibilità e la versatilità del laser e del sistema produttivo nel quale esso viene impiegato.

Ci si sta abituando a parlare di due categorie di laser ad alta potenza, leggeri e pesanti. La classificazione dipende principalmente dalla potenza. I laser leggeri sono caratterizzati da una gamma di potenza compresa fra qualche decina e qualche centinaio di watt. Vengono impiegati per i lavori di taglio e foratura dei supporti in ceramica e nell'industria elettronica, nella foratura dei rubini da orologio e per il taglio non solo dei metalli, ma anche di tessuti, materie plastiche e legno in molte altre industrie. Molti laser sono dispositivi allo stato solido di modeste dimensioni: i laser a rubino (con lunghezza d'onda di 0,69 micrometri), laser a vetro neodimio e laser ad alluminio ittrio neodimio (ambedue con lunghezza d'onda di 1,06 micrometri nell'infrarosso). Tali lunghezze d'onda si accoppiano bene alla maggior parte dei metalli rendendone possibile l'impiego per saldatura, foratura, taglio e trattamenti termici. I laser a rubino e vetro neodimio sono impiegati di solito in operazione impulsata con la quale si eroga una rapida scarica di impulsi di energia di alta intensità sulla superficie del pezzo da lavorare. I laser ad alluminio ittrio neodimio sono adatti per il funzionamento in onda continua (fascio costante), oppure in modo impulsato, con erogazione di parecchie migliaia di impulsi al secondo.

La categoria dei laser leggeri comprende anche alcuni laser a gas (argo e anidride carbonica) che vengono usati nella maggioranza dei casi in onda continua. Il fascio generato dal laser a gas presenta una collimazione quasi totale, cioè esibisce solo in lieve misura quella divergenza che caratterizza, per esempio, il fascio di una torcia elettrica. Pertanto il fascio può essere concentrato in un punto molto ristretto (dai micrometri a una frazione di millimetro) con erogazione di potenza ad alta intensità. Tali caratteristiche sono importanti, particolarmente nel caso della saldatura di profonda penetrazione che vedremo più avanti. I laser a gas sono anche molto efficienti; nel caso del laser ad anidride carbonica (CO_2) si può raggiungere il 15 per cento dalla linea alla potenza del fascio.

I laser cosiddetti pesanti vanno da pochi chilowatt a qualche decina di chilowatt. Tuttavia, attualmente vi sono solo dispositivi sperimentali o di laboratorio in grado di superare i 20 chilowatt. Nei processi di fabbricazione i laser pesanti vengono impiegati per lavorazioni metalmeccaniche pesanti come la saldatura di tubazioni per condotte, la saldatura di componenti automobilistici e il trattamento termico superficiale di alberi motori e delle pareti dei cilindri di grandi motori diesel. Il trattamento indurisce la superficie incrementando la resistenza all'usura dei particolari. La maggior parte dei laser per impieghi gravosi sono del tipo ad anidride carbonica con funzionamento in onda continua.

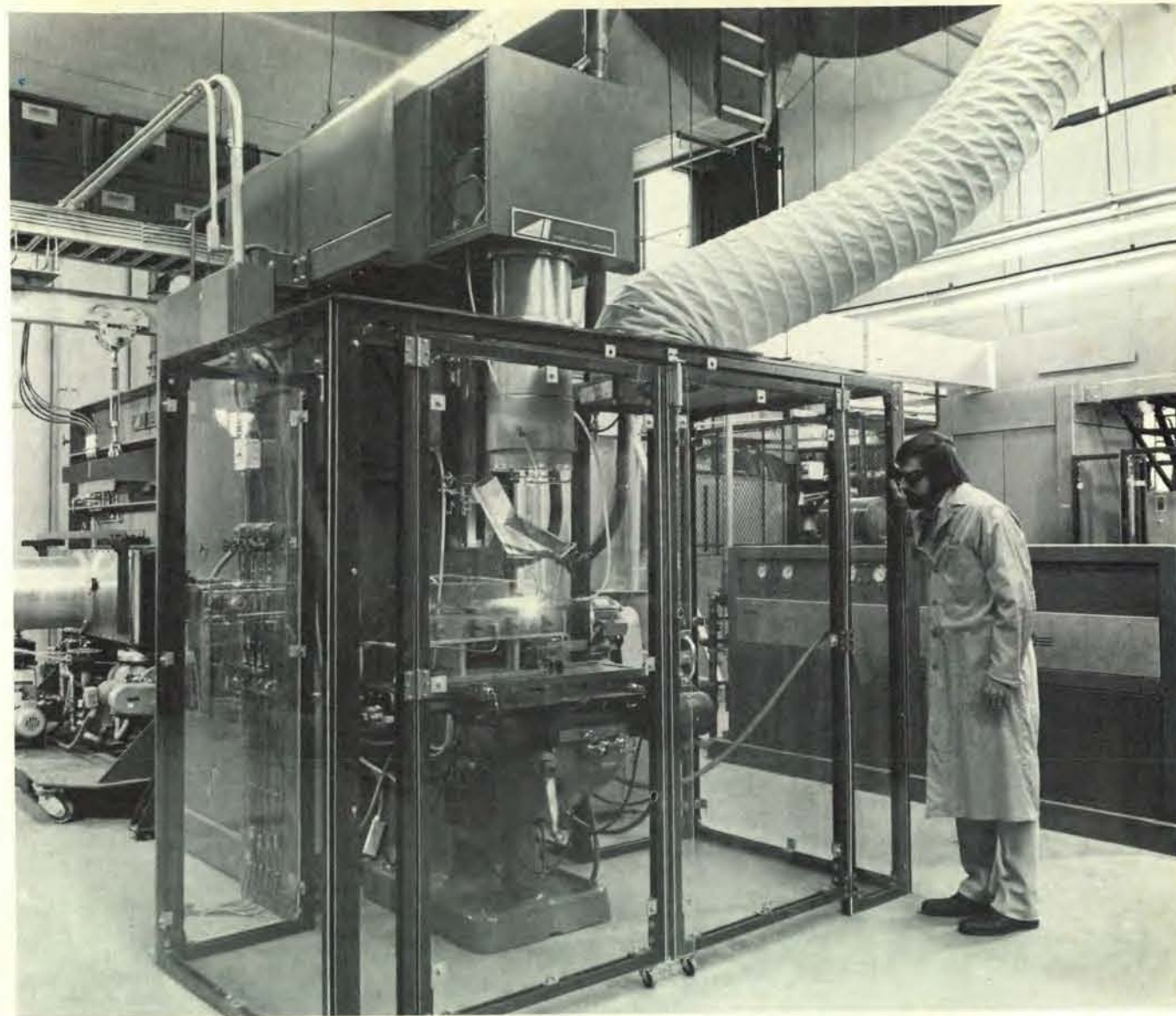
Il grande flusso di energia elettromagnetica applicato alla superficie del pezzo in lavorazione dal laser viene assorbito in corrispondenza di uno strato esterno di spessore pari a circa 10 nanometri (1 nanometro è pari a 10^{-6} millimetri). In tale strato sottile si crea una sorgente di calore di altissima intensità. Uno dei vantaggi del laser è che l'energia termica viene mantenuta e sfruttata nella zona in cui si deve effettuare la lavorazione. Per tale motivo il rendimento energetico risulta alto, da 10 a 1000 volte superiore a quello dei sistemi tradizionali che riscaldano volumi proporzionatamente maggiori del pezzo. Pertanto i sistemi laser fanno registrare tempi di lavorazione molto ridotti ed esibiscono proprietà uniche.

La saldatura laser di un anello sincronizzatore per cambio di autovettura, come viene eseguita presso lo stabilimento Fiat Mirafiori di Torino. Il fascio laser invisibile viene erogato da un ugello conico all'estremità della testa di saldatura appena sopra le scintille causate dall'ablazione esplosiva delle impurità superficiali durante una prima passata di imbastitura effettuata ruotando il pezzo una volta sotto il fascio alla velocità di 75 giri al minuto. Durante la seconda rotazione alla velocità di 25 giri al minuto in esecuzione della saldatura di profondità si verificano solo poche scintille. Il fascio è generato da un laser ad anidride carbonica dalla potenza di 2,5 chilowatt prodotto dalla Spectra-Physics.

Rappresentazione schematica di un laser industriale da 1200 watt a radiazione infrarossa continuativa del tipo a trasporto di gas, prodotto dalla Spectra-Physics. Una miscela di anidride carbonica, azoto e elio, viene portata ad un alto livello energetico in non equilibrio termodinamico (condizione necessaria per la generazione di luce laser) mediante una scarica elettrica tra catodo e anodo. Il fascio laser viene amplificato durante sette passaggi tra gli specchi del sistema ottico: si riscalda e viene quindi inviato allo scambiatore di calore. Il trasporto di gas avviene in un anello continuo in cui il gas, nell'ordine, è eccitato, contribuisce al fascio laser e viene raffreddato.

Questa raffigurazione in prospettiva della unità laser da 1200 watt mostra il ciclaggio del gas e il piegamento del fascio. Quando emerge dallo specchio di uscita il fascio viene portato in corrispondenza della testa di lavoro per la saldatura, il taglio, la foratura o il trattamento termico.

Un altro importante vantaggio del laser è quello di non danneggiare i pezzi in lavorazione in quanto esso eroga calore in minor tempo delle sorgenti tradizionali a causa dell'alta densità di potenza del fascio. Pertanto il calore non ha tempo di propagarsi al pezzo. Le sorgenti tradizionali riscaldano il pezzo molto più del necessario con conseguenti deformazioni, incrinature o sollecitazioni che possono essere causa di danni e che rendono necessaria l'ulteriore lavorazione o la riduzione in rottami o che più semplicemente ne compromettono le prestazioni. Le conseguenze economiche sono ovvie nel caso di particolari semilavorati costosi come gli ingranaggi i cui denti devono essere induriti, le pale delle turbine per motori a reazione sulle quali devono essere praticati i fori di raffreddamento e i blocchi motore i cui cilindri devono essere induriti.

Tali vantaggi sono resi possibili dalla densità di potenza estremamente alta del fascio laser. Taluni altri vantaggi rendono il fascio laser uno strumento molto flessibile e spiegano il motivo per cui i laser vengono spesso impiegati con buoni risultati anche ai livelli di densità di potenza ottenibili mediante le sorgenti tradizionali. Il fascio non possiede massa, quindi può essere facilmente spostato e controllato in tempi di risposta brevi. Si presta a essere inserito in cicli di lavorazione automatici con estrema facilità. Agisce a distanza eliminando o riducendo i problemi di interferenza meccanica. Non genera allo stesso tempo reazioni di natura meccanica e quindi il pezzo non vibra e non necessita di bloccaggio alcuno. Inoltre, la tecnologia laser permette una lavorazione pulita e veloce altamente compatibile con le unità operatrici meccaniche distribuite lungo la linea di produzione. Pertanto essa ha importanti implicazioni in termini della logistica e della compattezza del sistema produttivo.

Tali vantaggi diventano ovvi quando si parla delle applicazioni specifiche del laser nei processi lavorativi. Enunceremo tali applicazioni nell'ordine in cui esse sono state accettate nella industria e cioè foratura, sbavatura, taglio, saldatura e trattamenti termici.

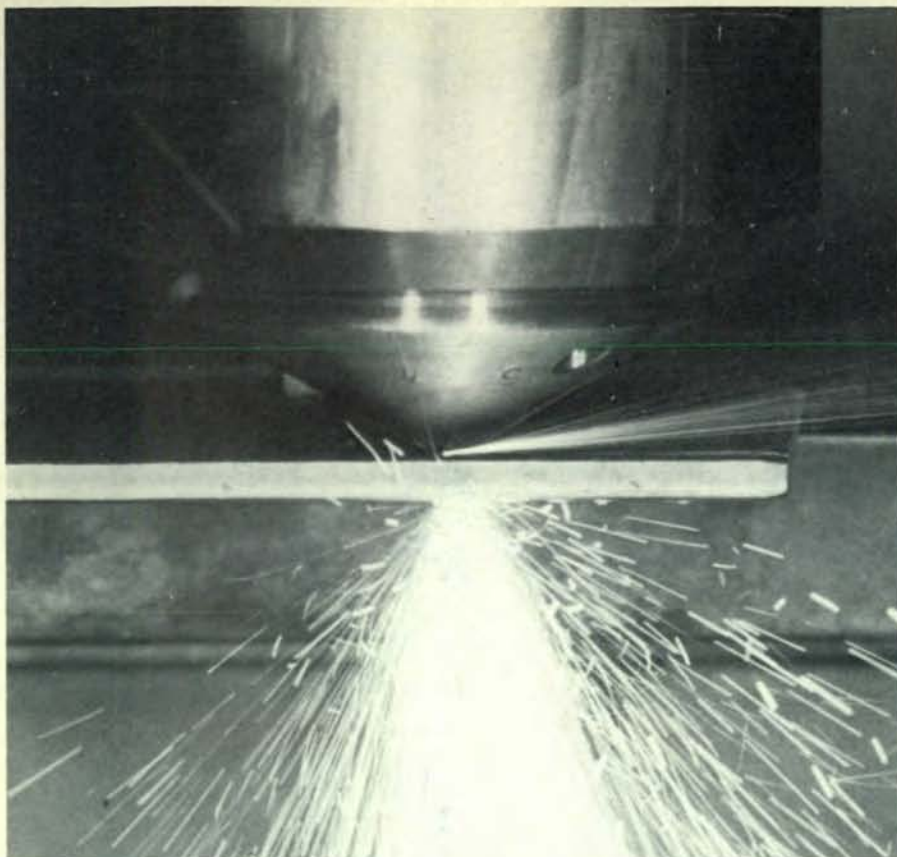
Nella foratura e nella sbavatura la funzione del laser è quella di asportare il materiale. La foratura mediante laser è un processo energeticamente molto costoso in quanto il materiale deve essere prima fuso e poi vaporizzato (la rimozione della fase liquida risulta difficile per fori di profondità superiore a due volte il diametro). Infatti, i calori latenti di vaporizzazione sono per la maggioranza dei metalli circa 10 volte maggiori di quelli di fusione.

Va aggiunto che il vapore generato e surriscaldato risulta altamente suscettibile alla eccitazione e ionizzazione dal fascio laser e che le corrispondenti energie di trasformazione nella fase ionizzata sono di almeno 100 volte superiori a quelle coinvolte nella fusione della maggior parte dei metalli. Il vapore ionizzato si

Nella fotografia è visibile un sistema di saldatura e taglio per il trattamento termico di superfici metalliche che impiega un laser Avco Everett da 15 chilowatt della prima generazione. La stazione di lavoro sul fondo a destra è adibita al trattamento delle superfici; quella visibile in primo piano è adibita alle operazioni di saldatura e taglio. L'operazione in fase di svolgimento consiste nella saldatura di un pezzo di acciaio inox da mezzo pollice di spessore alla velocità di 25 pollici al minuto con una potenza di fascio pari a 7,5 chilowatt. Il laser eroga un fascio infrarosso continuo che viene diretto verso il basso sulla superficie del pezzo tramite un telescopio. Le traslazioni di lavoro avvengono mediante una normale tavola di fresatrice. La potenza del fascio può essere regolata, a seconda delle esigenze di lavoro, da 1,5 a 15 chilowatt.

combina con i gas circostanti per formare una «piuma» di plasma: un gas ionizzato. La piuma di plasma assorbe, riflette e disperde la radiazione laser impedendole di raggiungere il pezzo. Il fenomeno è noto come *blanketing*. In genere, viene controllato sostituendo il gas ambientale con un gas inerte tipo l'argo o l'elio oppure una miscela dei due. Essi sono maggiormente trasparenti alla radiazione e meno facilmente eccitabili e ionizzabili.

Per la foratura si è trovato che operando alle maggiori densità di potenza, normalmente ottenibili mediante l'impiego di laser impulsati, ci si può avvalere di una serie di fenomeni plasma-fluidodinamici, associati alla formazione della piuma. Per loro tramite è possibile innescare dei processi simili alla combustione sostenuta dal laser e, con l'incremento dell'intensità di potenza, a una detonazione sostenuta dal laser. Le pressioni e le onde d'urto del plasma in uscita dal foro ad alta velocità (supersonica con detonazione) risultano molto efficaci nel rimuovere il materiale fluido e gli eventuali prodotti solidi come scorie, ossidi o nitruri. Pertanto la foratura viene eseguita prevalentemente mediante laser allo stato solido in quanto meglio controllabili in operazione impulsata che i laser a gas.

Tutto sommato la foratura mediante laser rimane un procedimento costoso. Risulta pertanto giustificata su scala industriale soltanto per fori di piccole dimensioni in materiali di difficile lavorazione con mezzi meccanici a causa della loro tendenza a impaccare gli utensili, della loro durezza o della loro abrasività, oppure perché sono così teneri per cui non è possibile praticare fori di alta precisione. Comunque, per la lavorazione di tali materiali, i vantaggi del laser sono così netti che quest'ultimo ha quasi completamente sostituito i metodi di foratura tradizionali.

Il più grande successo è stato quello della foratura di rubini per l'industria degli orologi. Il laser viene anche largamente usato nella foratura dei diamanti impiegati per matrici di trafilatura per fili metallici. Nell'industria dell'elettronica il laser viene impiegato per la foratura e il taglio dei supporti ceramici per circuiti integrati e per la rifilatura di tali circuiti e dei componenti elettronici discreti. L'industria aerospaziale impiega il laser per praticare fori su componenti di materiali eccezionalmente duri come quelli delle pale per turbine e dei reggifiamma delle camere di combustione. All'altro estremo

Questo taglio laser viene eseguito presso l'Istituto di tecnologia meccanica di Vico Canavese. Una piastra di acciaio inox di 5 millimetri di spessore viene alimentata da sinistra verso destra alla velocità di 1,2 metri al minuto al di sotto della testa di taglio fissa. Il fascio laser a raggi infrarossi invisibili passa attraverso un ugello concentrico per l'erogazione dell'ossigeno che reagisce con il metallo riscaldato dando origine alla combustione che accelera l'operazione. Le scintille sono causate dall'espulsione di particelle liquide o solide dalla zona di taglio. L'operazione è effettuata a 2,5 chilowatt mediante un fascio generato da un laser Avco programmabile da 1,5 a 15 chilowatt.

della scala delle durezze, il laser trova sempre più numerose applicazioni per la foratura in materiali morbidi e flessibili come quelli impiegati nella fabbricazione di tubi e cavi (gomma e polietilene), nella produzione delle valvole in plastica per i contenitori di aerosol e per i capezzoli di gomma per poppatoi per lattanti.

Il fattore comune a tutte queste applicazioni è che i fori da praticare sono di piccole dimensioni. Il costo energetico di foratura mediante laser risulta pertanto inferiore. In tutte queste applicazioni i laser stanno rapidamente sostituendo i mezzi convenzionali di lavorazione e, malgrado il loro costo iniziale molto maggiore, viene asserito che si ottengono tempi d'ammortamento estremamente brevi a causa della loro grande produttività e dell'eliminazione di utensili costosi per acquisto, manutenzione e stoccaggio.

Nelle industrie metalmeccaniche pesanti le forature con il laser risultano troppo costose a causa della grande quantità di materiale da asportare. Per queste industrie è di qualche interesse l'impiego del laser per le operazioni di foratura finale, in genere con fori di minor diametro, impiegate per aprire fori non passanti e per l'operazione finale di sbavatura e (oppure) per l'eliminazione dei trucioli quando queste operazioni diventano difficoltose per ragioni di accessibilità o per evitare di danneggiare le finiture delle superfici in cui i fori devono venire aperti. Cilindri freno, iniettori diesel, alberi motore e simili pezzi sembrano attualmente quelli che più facilmente potranno sfruttare il processo laser in produzione industriale.

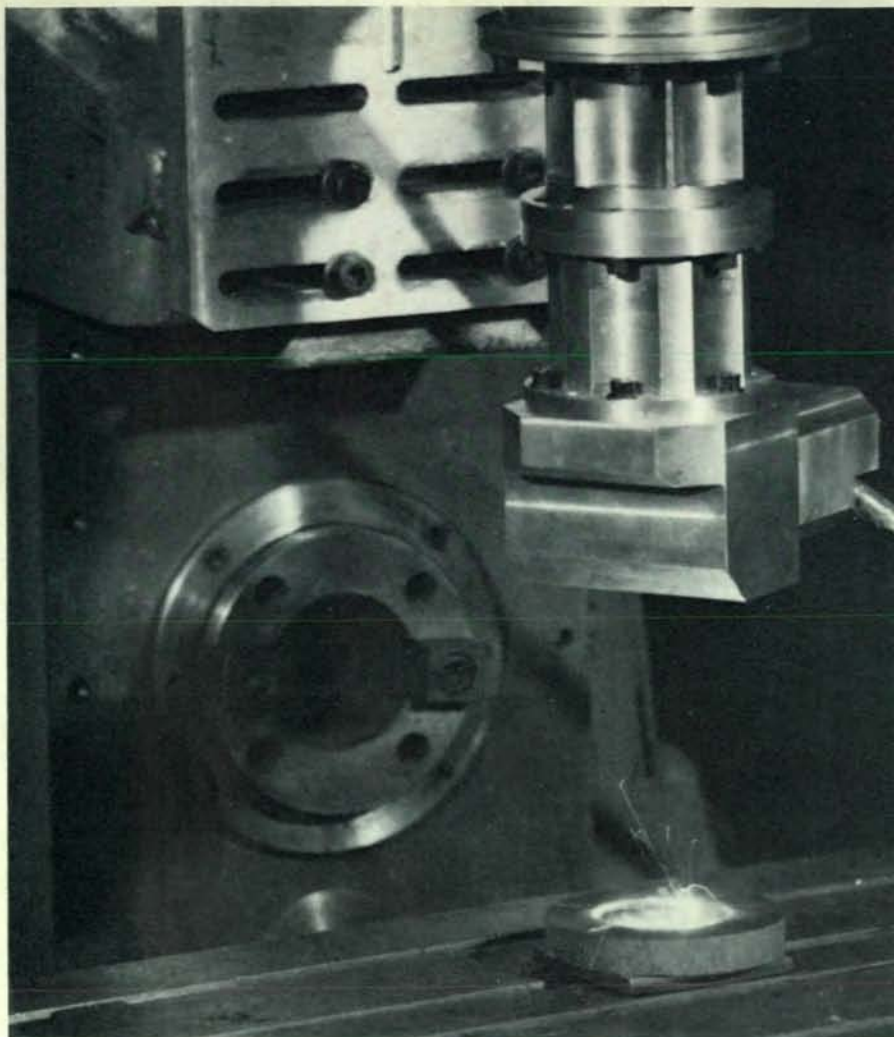
Paragonato alla foratura, il taglio consuma minore energia per volume di materiale asportato poiché richiede che il materiale venga portato solo alle condizioni di fusione basandosi per la rimozione della fase liquida sull'azione meccanica di un getto di gas di copertura, generalmente un gas nobile che viene anche usato per proteggere il pezzo e le superfici tagliate da indesiderabili effetti di ossidazione e nitrurazione. Quando questi non rappresentano un problema si possono usare altri gas, come per esempio azoto e ossigeno. Quando l'ossidazione della superficie non dà problemi, l'ossigeno può essere usato per incrementare più volte la velocità di taglio a causa delle sue reazioni esotermiche con la maggior parte dei metalli alle alte temperature. Gli acciai ad alto tenore di carbonio e le leghe di titanio sono buoni esempi poiché i prodotti di reazione sono in fase solida, facilmente asportabili dal getto di gas.

Il fascio laser si presta a essere facilmente controllato e automatizzato e quindi con l'impiego di elaboratori a controllo numerico può essere diretto a seguire le forme più complesse, iniziando indifferentemente dai bordi o dall'interno del foglio di materiale, prestazioni che sono difficili se non del tutto impossibili con certi mezzi meccanici, per esempio ruote diamantate, spesso le sole capaci di attaccare questi materiali. Per questo motivo sembra probabile che l'impiego del laser riuscirà a estendere l'utilizzazione di tali materiali (riducendo sfridi e consentendo il taglio in forme altrimenti impossibili).

La programmazione automatica del taglio laser comporta grandi miglioramenti in precisione, ripetibilità, flessibilità e produttività. La sinergia di questi fatti sull'intero ciclo lavorativo della fabbrica risulta in incredibili aumenti in produttività che si valutano talvolta da otto a 20 volte. Altri vantaggi sono: mancanza di vibrazioni, di rumori, di polveri e fumi e il non dover più acquistare, stoccare e mantenere una varietà e quantità di utensili spesso molto costosi.

Questo spiega perché il laser, malgrado il suo costo iniziale, che può essere maggiore, talvolta anche più di 5 volte, del costo dei sistemi convenzionali, riesca a dimostrarsi molto attraente anche per ragioni economiche. I sistemi laser permettono tempi di ammortamento estremamente brevi (meno di un anno) e questo viene anche indicato dalla sua rapida e crescente velocità di espansione in una grande varietà di lavorazioni di taglio. Si può stimare che su di una scala mondiale si sia già nell'ordine di ben più di 100 sistemi per anno nelle potenze da 400 a 2000 watt.

Fra le industrie in cui queste applicazioni sono più diffuse si possono indicare: quelle della lavorazione delle lamiere metalliche prerivestite, materiali di fibra organica (legno, cuoio, cartone e tessili), materiali plastici con e senza rinforzo di fibra, gomma per tubazioni flessibili rinforzate con filo metallico, materiali ceramici, quarzo, vetro, materiali compositi ed esotici per impieghi aerospaziali. Il laser offre vantaggi inattesi anche per i materiali facili da tagliare, come i tessuti per la confezione degli abiti.

La lavorazione di tali tessuti richiede spesso il taglio di modelli complessi. Inoltre, il taglio deve essere preciso e netto, senza sfilacciature onde ovviare alle operazioni di cucitura dei bordi. Sistemi laser automatici azionati a mezzo di calcolatore numerico hanno avuto notevole successo in queste applicazioni. Il tessuto viene tagliato a uno strato singolo per volta, eliminando la tendenza all'incollamento degli strati multipli nella zona di taglio. L'alta velocità di lavorazione rende possibile il taglio di tessuto sufficiente per oltre 40 abiti per ora mediante laser di potenza relativamente bassa (da 200 a 400 watt).

Questo allegamento di sede valvola mediante laser è eseguito presso il Centro ricerche Fiat di Orbassano. Il calore intenso erogato dal fascio laser in una zona concentrata rende possibile l'allegamento superficiale del pezzo dopo la formatura. Questa sede valvola in ghisa viene allegata con polvere di elementi scelti in sostituzione della stellite, una lega contenente circa il 50 per cento di cobalto, cioè un materiale strategico molto costoso. La nuova lega viene sottoposta a prova per misurarne le prestazioni rispetto alla stellite per quanto concerne la resistenza all'usura in presenza di gas di scarico caldi e chimicamente reattivi nella zona in cui la testa della valvola farà ripetutamente contatto. Durante questo trattamento termico il fascio laser provoca la fusione e il mescolamento degli elementi leganti che di conseguenza si diffondono rapidamente nella ghisa.

Questo ingrandimento della sede valvola mostra il punto di erogazione del calore ad alta intensità da parte del fascio laser con spostamento automatico dello stesso lungo la superficie di allegamento. I materiali da mescolare alla ghisa vengono applicati sotto forma di polvere. Poiché il fascio laser non stabilisce contatti meccanici con il pezzo, la sede valvola non richiede bloccaggio.

Uno di tali sistemi, realizzato dal Culham Laboratory in Gran Bretagna, consiste in un laser da 400 watt il cui fascio è guidato mediante specchi mobili. Il fascio effettua il taglio del tessuto in movimento di due metri di larghezza a una velocità di avanzamento che può raggiungere gli 80 metri al minuto. La memoria del calcolatore contiene tutte le informazioni necessarie concernenti gli stili, le taglie e le varianti relative a ordinazioni speciali. Una ventina di tali sistemi sono in produzione negli Stati Uniti e in Gran Bretagna.

Due tipi di saldatura possono essere eseguiti con il laser: quella per conduzione e quella di profonda penetrazione. Nella saldatura per conduzione la sorgente di calore viene mantenuta alla superficie del pezzo. La superficie deve essere mantenuta a una temperatura inferiore a quella di vaporizzazione per evitare la perdita di materiale che comporterebbe l'indebolimento della saldatura. Le saldature così ottenute sono di poca profondità (rapporto di figura 1:1,5) il che significa che la profondità della saldatura può solo essere 1,5 volte la larghezza. Pertanto, la saldatura per conduzione serve principalmente solo per la giunzione di lamiere o piastre di piccolo spessore. Non è molto efficiente (dall'8 al 15 per cento) e lascia una zona termicamente alterata di dimensioni piuttosto ampie.

Di maggiore interesse è la saldatura di profonda penetrazione. In questo caso, accettando di vaporizzare parte del materiale, la sorgente di calore può essere portata giù dalla superficie all'interno del materiale. Il rendimento del processo aumenta marcatamente a valori di oltre l'80 per cento anche in presenza di fenomeni energeticamente dispendiosi come la vaporizzazione, l'eccitazione, la ionizzazione e il blanketing.

Una spiegazione di tale fatto è data dagli effetti benefici generati dall'interazione di molti fenomeni complessi che si verificano allorché la sorgente di calore penetra in profondità nel pezzo e crea una zona con un cuore di vapore surriscaldato circondato da una zona di materiale fuso anch'esso surriscaldato. Il cuore occupa una piccola parte che nella saldatura a piena penetrazione si estende a tutto lo spessore del materiale. Tale zona viene denominata buco guida *(keyhole)*. Un fascio di alta intensità correttamente a fuoco può produrre un effetto «keyhole» in pochi millisecondi per poi mantenerlo in equilibrio dinamico mentre lo sposta man mano che si forma la saldatura.

Il materiale fuso viene sollevato al di sopra della superficie superiore dalla pressione del vapore surriscaldato e trattenuto da una combinazione di forze di gravità, viscosità e tensione superficiale, che prevengono anche la sua caduta dalla superficie inferiore del «keyhole». Di conseguenza le proprietà meccaniche della saldatura risultano migliorate poiché, senza apporto di materiale, essa presenta cordoni alle due superfici. Si deve notare che a causa dei gradienti di temperatura e di pressione il materiale fuso viene violen-

temente mescolato nella zona che circonda la sorgente di calore e, quindi, si ha un trasferimento di calore a causa sia del trasporto di massa sia del calore latente dovuto alle trasformazioni di fase (come per esempio quella da gas a liquido). Pertanto il «keyhole» vien fatto agire come un tubo di calore, e per giunta come uno di ben maggiore efficacia di ogni altro tubo di calore, in quanto sfrutta il calore latente dell'eccitazione e della ionizzazione che sono di ordini di grandezza superiori a quelli delle trasformazioni di fase usate nei comuni tubi di calore.

L'altissima capacità di trasporto del calore stabilitosi nel «keyhole» significa che il materiale nella zona circostante agisce da isolante impedendo il flusso non desiderato di calore alla parte restante del pezzo. Questo fenomeno spiega l'alto rendimento della saldatura laser a profonda penetrazione che è anche evidenziato dall'alto rapporto di figura (da 1:5 a 1:12) e da una zona termicamente alterata molto contenuta ai lati della saldatura.

Per quanto concerne l'aspetto metallografico, la saldatura di profonda penetrazione ottenuta mediante il fascio laser presenta una buona struttura cristallina con poche impurità. Di conseguenza, le proprietà meccaniche come la durezza, la resistenza a trazione e la resilienza risultano uguali a quelle del materiale base e in alcuni casi anche migliori. Le azioni risultanti di affinamento e di purificazione sono evidenti su materiali quali l'acciaio per costruzioni navali e sul nuovo acciaio basso-legato ad alta resistenza.

Inizialmente la saldatura laser dell'alluminio e delle sue leghe ha incontrato alcune difficoltà a causa della bassa tensione superficiale e della bassa viscosità della zona fusa. Sulle lamiere di spessore superiore ai tre millimetri la zona fusa risultava instabile e, quindi, la zona di saldatura era affetta da depressioni e porosità. Nel corso di un programma di ricerche sponsorizzato dalla NASA tali problemi vennero superati per le piastre di lega di titanio e alluminio mediante il preriscaldamento. Infatti, attualmente la saldatura di piastre da mezzo pollice non presenta più alcun problema.

La saldatura mediante laser offre molti vantaggi pratici. Non richiede l'uso di elettrodi, flussi o materiali protettivi. È possibile saldare materiali dissimili. La possibilità di deformazioni, tensioni interne e cricche viene ridotta al minimo il che è particolarmente importante nel caso di materiali da saldare dissimili o fragili. Non è necessario stabilire contatto fisico con le parti da saldare: è sufficiente che esse siano raggiungibili dal fascio. Il fascio può essere guidato mediante specchi il che significa che lo si può far impattare all'interno di particolari cavi e in zone generalmente poco accessibili senza introdurre alcuna sollecitazione meccanica. Per esempio, con l'impiego di specchi i sistemi laser si sono dimostrati capaci di effettuare saldature di profonda penetrazione all'interno di condutture.

La flessibilità e la facilità di movimentazione, posizionamento e controllo del fascio laser ne facilitano grandemente l'automatizzazione, e, quindi, si possono realizzare sistemi altamente produttivi per la lavorazione di strutture e forme complesse. In tali applicazioni la tecnologia laser riduce o elimina del tutto la necessità di personale altamente specializzato semplificando molto l'arte, finora ritenuta «magica», di eseguire saldature difficili.

D'altra parte, la saldatura mediante laser richiede maggiore precisione di posizionamento dei pezzi. Se la distanza tra le parti da saldare è superiore al 25 per cento dello spessore dei particolari stessi e non viene impiegato materiale di apporto, si può verificare la caduta del materiale base con conseguenti saldature di insufficiente spessore (depresse) o porose. Tuttavia, la precisione di posizionamento non ha costituito un serio problema.

La foratura mediante laser permette di realizzare complesse figure di minuscoli fori di raffreddamento nella paletta di un motore a reazione. Per il buon funzionamento del particolare è indispensabile tenere sotto stretto e preciso controllo il diametro e la profondità dei forellini e l'angolo di entrata del fascio laser. Questa operazione di foratura viene effettuata presso la Pratt & Whitney Aircraft Division della United Technologies Corporation con un laser a rubino che eroga 45 chilowatt di potenza in meno di un millisecondo per impulso. Il laser per applicazioni industriali per realizzare i fori di raffreddamento è stato prodotto dalla United Technologies.

Le applicazioni industriali delle saldature laser si riscontrano nelle industrie elettriche ed elettroniche nel caso di materiali dissimili, di materiali prerivestiti, di fili ai loro terminali, di elettrodi per batterie, per l'impacchettamento di laminazioni per motori elettrici. Come sempre, la saldatura laser è migliore di quella convenzionale in quanto richiede una minore quantità complessiva di calore e, quindi, la zona termicamente alterata risulta contenuta. L'affinamento metallurgico possibile con la saldatura laser di materiali dissimili evita la possibilità di infragilimento del giunto. Nel caso dei materiali prerivestiti, il fascio laser asporta il rivestimento della zona da saldare simultaneamente all'esecuzione della saldatura mentre con i metodi convenzionali le operazioni devono essere eseguite separatamente. Nella saldatura degli elettrodi alle batterie e dei fili ai terminali il contenimento della zona termicamente alterata significa che il calore di saldatura non arreca danni alle parti circostanti della batteria o del circuito.

Un esempio di ciò che si può realizzare con la saldatura laser è dimostrato dall'impianto automatizzato della Hamilton Standard Division (della United Technologies Corporation) per la saldatura degli elettrodi di batterie. L'apparecchiatura si avvale di un laser ad anidride carbonica da tre chilowatt di produzione della United Technologies. Vien detto che l'impianto, impiegato in tre stabilimenti diversi, ha fatto registrare un incremento della produttività da sette a 10 volte superiore a quella dei sistemi convenzionali offrendo al tempo stesso un controllo di qualità di gran lunga superiore.

Le applicazioni più pesanti che comportano saldature di particolari di spessore oltre i quattro millimetri sono allo stato di transizione tra la sperimentazione e la realizzazione. Diverse aziende stanno sperimentando il laser per la saldatura di

strutture navali, tubazioni, ponti e prefabbricati in acciaio.

Le industrie aerospaziali stanno adottando la saldatura laser in sostituzione di quella a fascio elettronico in un certo numero di applicazioni a causa dei vantaggi offerti da un'operazione fuori dal vuoto, in applicazioni come la saldatura di costole di rinforzo all'interno di fusoliere di missili, la saldatura interna di armadiature per strumentazioni, la saldatura di serbatoi per gas ad alta pressione costruiti con alluminio e titanio e la saldatura di ingranaggi e palette per motori a turbina. Le industrie automobilistiche sono nelle prime fasi dell'applicazione industriale del laser alla saldatura di precisione di pezzi prodotti in grandi serie come scatole frizione e differenziale, ingranaggi, assali, sensori dei gas di scarico e componenti per impianti frenanti.

L'applicazione del calore generato dal laser per il trattamento superficiale dei particolari è un processo che riveste grande importanza per molte industrie di produzione in grandi serie. Lo scopo di queste operazioni, indipendentemente dalla sorgente di calore, è di dare alla superficie del pezzo certe proprietà speciali necessarie per un funzionamento affidabile di lunga durata in particolari severe condizioni di funzionamento, come ad esempio la resistenza all'usura in presenza di alti carichi meccanici combinati a un'alta temperatura e in ambienti chimicamente aggressivi. Due dei principali tipi di trattamenti termici sono il riporto e l'allegamento superficiale.

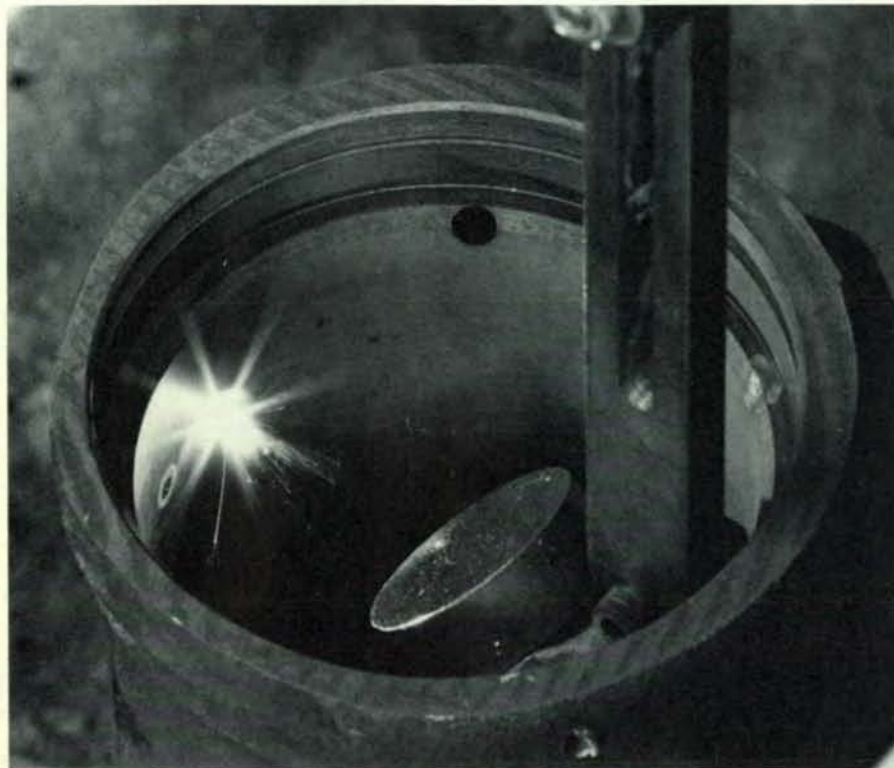
Il riporto aggiunge sul materiale di base del pezzo uno strato di materiale che possiede le proprietà desiderate. Un'applicazione tipica è il riporto di una lega di cobalto su acciaio. Il laser introduce in questo processo una più alta produttività, un'operazione pulita e controllabile con sistemi lavorativi che possono essere facilmente e completamente automatizzati e programmati. Il riporto può ora essere effettuato sulle linee di produzione meccaniche e non è più relegato a officine metallurgiche fuori linea.

L'allegamento superficiale è un processo che il laser ha reso industrialmente possibile. I materiali da aggiungersi possono essere polveri, bacchette, fogli o anelli delle leghe, oppure le polveri degli elementi stessi che possono essere allegati in loco. Le implicazioni per la progettazione e la produzione dei pezzi sono notevoli e abbastanza ovvie. Nel processo di allegamento vengono impiegate temperature più alte. Esse sono raggiunte rapidamente operando con maggiori intensità di potenza, creando elevati gradienti di temperatura a mezzo dei quali il materiale aggiunto viene fuso, mescolato violentemente e fatto diffondere nel sottostante materiale di base che può stare ancora allo stato plastico. A causa dei valori molto elevati delle forze che attivano il processo, si ottengono velocità di diffusione molto grandi. Poiché le velocità di raffreddamento sono molto alte, le leghe così generate posseggono una microstruttura molto fine e la distribuzione degli elementi leganti può essere opportunamente regolata mediante appropriate selezioni delle modalità con cui il campo termico generato dal laser viene applicato.

L'allegamento superficiale mediante laser acquista grande importanza per le insolite caratteristiche che conferisce ai particolari e l'alta qualità dei prodotti, a cui si sommano importanti vantaggi economici rappresentati da una sostanziale riduzione dei consumi di materiali costosi in quanto il laser permette di applicare tali materiali solo nelle zone in cui si richiedono alte prestazioni. Ulteriori vantaggi sono dovuti al fatto che gran parte del materiale risulta più facilmente lavorabile e solo una minima entità di lega dura deve essere asportata con le operazioni di finitura (che richiedono l'impiego di utensili costosi come le mole diamantate in quanto sono le sole in grado di attaccare alcune di queste leghe).

Un terzo tipo di trattamento termico è la trasformazione di fase in base alla quale una superficie può essere indurita (senza riporto di materiale) provocando il passaggio del materiale base da una fase solida a un'altra di maggior durezza. È importante, per evitare l'infragilimento del pezzo, che l'indurimento non penetri troppo in profondità. In genere è sufficiente produrre uno strato indurito di pochi millimetri di spessore per assicurare un adeguato supporto alla superficie resistente. Lo stesso spessore sembra essere sufficiente a creare uno stato tensionale di compressione capace di migliorare la resistenza a fatica del pezzo.

Nell'operazione di tempra superficiale si fa in modo che il profilo di temperatura, generato a mezzo dell'irradiazione laser, sia tale da mantenere la superficie sotto il suo punto di fusione mentre fino a una profondità desiderata vengano raggiunte le temperature necessarie per ottenere le volute transizioni di fase. Poiché il calore viene applicato solo a un sottile strato superficiale, ne consegue un rapido processo di autoraffreddamento causato dalla massa sottostante del materiale. A causa dei tempi molto brevi di raffreddamento si ottengono microstrutture molto fini che al limite, nel cosiddetto *glazing process*, possono risultare in uno strato totalmente amorfo o vetroso.

La tempra mediante trasformazione di fase rappresenta attualmente l'unico tipo di trattamento termico effettuato mediante laser su scala industriale. Una importante applicazione si trova alla General Motors Corporation, presso la Saginaw Steering Gear Division, dove oltre venti sistemi laser vengono impiegati per l'indurimento delle superfici di guida nella scatola dei servosterzi ivi prodotti in quantità di 30 000 pezzi al giorno. Ancora la General Motors ha in produzione presso la sua Electromotive Diesel Division diversi sistemi laser da cinque chilowatt per l'indurimento delle superfici interne delle canne cilindri in ghisa

La tempra superficiale interna della scatola servosterzo mediante laser è eseguita presso la Saginaw Steering Gear Division della General Motors Corporation. La scatola è di ghisa malleabile ferritica. Il fascio laser all'anidride carbonica di 1000 watt passa inferiormente attraverso una lente focalizzatrice e arriva a uno specchio con placcatura in oro che si sposta verticalmente all'interno della scatola per effettuare la tempra superficiale in strisce equidistanti. Il fascio laser è invisibile: il punto luminoso della fotografia indica la zona della scatola riscaldata dal fascio.

per grandi motori diesel prodotti in quello stabilimento.

Sebbene i successi descritti siano notevoli, essi sono solo gli inizi di un contributo ben maggiore che i sistemi laser sono in grado di apportare all'industria. Quelli attuali sono sistemi laser della prima generazione di potenza limitata (da 200 a 2000 watt). I laser di maggiori potenze stanno diventando disponibili e vengono adottati in numero sempre maggiore sulle linee di produzione. È da prevedersi che quando sistemi di provata affidabilità nel campo delle potenze da 10 a 20 chilowatt diventeranno disponibili, essi avranno una grande influenza su di un settore molto più importante per le economie dei paesi industrializzati, ovvero sul settore delle industrie di produzione di massa. Perché questa transizione avvenga e in maniera efficace, occorre che le nazioni e le industrie interessate provvedano a prepararsi. Questo implica non solo lo sviluppo di sorgenti laser più avanzate, ma uno sforzo globale che riveda l'intero spettro dei processi industriali per assicurarsi che tutto il potenziale di questa nuova tecnologia in produttività e versatilità venga appropriatamente usato.

La novità di questo nuovo mezzo che proviene da discipline e tecnologie con cui la presente generazione degli ingegneri e tecnici di produzione è scarsamente familiare, renderà inizialmente più difficile la sua introduzione nelle linee di produzione convenzionali. Pertanto sarà necessario usare un approccio pragmatico cominciando con applicazioni relativamente semplici e di sicuro successo. Molto probabilmente saranno monoscopo e punteranno per il loro successo sull'effetto combinato di diversi fattori favorevoli, piuttosto che sullo spingere a fondo la capacità produttiva del sistema laser.

Nello stesso tempo si dovrà iniziare la preparazione per una seconda fase in cui sistemi multiscopo e versatili saranno capaci di sfruttare completamente la produttività, resa possibile dal nuovo mezzo. Va detto che benché gli incrementi in produttività già ottenuti siano considerati sostanziali, essi non rappresentano che una piccola frazione di quelli possibili. Essi appaiono soddisfacenti solo perché il sistema produttivo odierno non è capace di accettarne maggiori, specie se derivanti da quelle lavorazioni di «tipo metallurgico», che sono proprio quelle a cui il laser ha più da offrire.

La situazione muterà radicalmente allorché il prodotto e gli impianti saranno progettati come un sistema integrato che sfrutti a pieno i vantaggi ottenibili con la tecnologia laser.

Si configureranno diverse linee di lavorazione multiscopo ad alta produttività in combinazione bilanciata facenti capo a un unico sistema laser di seconda generazione, centralizzato e altamente flessibile e in grado di effettuare in modo superiore tutte le operazioni di «tipo metallurgico» per le linee facenti capo a esso. Allora la produttività del laser, talmente alta da sembrare inutile per gli impianti produttivi oggi esistenti, potrà essere adeguatamente sfruttata e l'impatto rivoluzionario della tecnologia laser sarà pienamente sentito.

Superammassi e vuoti nella distribuzione delle galassie

Rilevamenti sistematici degli spostamenti verso il rosso in particolari regioni del cielo ci hanno rivelato che esistono almeno tre superammassi di galassie, ma anche che volumi immensi di spazio sono del tutto vuoti

di Stephen A. Gregory e Laird A. Thompson

Astronomi e cosmologi sono oggi molto impegnati nel tentativo di spiegare l'emergere e la distribuzione di aggregati di materia nell'universo. Quanto tempo dopo il «big bang», la presunta esplosione dell'atomo primordiale avvenuta fra 10 e 20 miliardi di anni or sono, la materia cominciò ad aggregarsi dando origine alle stelle e alle galassie che vediamo oggi? Supponendo che prima di quell'epoca la materia fosse dispersa in modo più o meno uniforme, l'universo su grande scala è oggi popolato in modo uniforme da sistemi stellari di qualche tipo? Osservazioni recenti compiute da vari gruppi di astronomi consentono oggi di cominciare a dare una risposta a queste domande. Rilevamenti sistematici su vasta scala hanno verificato l'esistenza di superammassi di galassie: strutture organizzate composte da molteplici ammassi di galassie. Ciascun ammasso, a sua volta, può essere composto da centinaia o migliaia di singole galassie. Benché l'esistenza di superammassi sia da molto tempo oggetto di congetture, la loro conferma si è accompagnata ad almeno una grande sorpresa: regioni altrettanto grandi dello spazio sono del tutto prive di galassie.

I superammassi hanno una vastità così sterminata che le singole galassie che ne fanno parte, che si muovono a velocità casuali, non possono aver percorso più di una frazione del diametro di un superammasso nei miliardi di anni trascorsi da quando hanno avuto origine. È perciò evidente che i superammassi possono permettere una comprensione della storia evolutiva dell'universo che risulta invece inaccessibile quando si limiti lo studio a sistemi minori. A scale inferiori a quelle dei superammassi la distribuzione originaria della materia è resa confusa dalla «mescolanza» evolutiva. Gli astronomi sperano che una comprensione delle strutture più grandi dell'universo possa aiutare a chiarire i processi che danno origine a strutture di ogni ordine di grandezza, dalle galassie alle stelle e ai pianeti.

È impossibile accertare chi sia stato il primo a concepire l'idea che gli ammassi di galassie potessero essere a loro volta membri di aggregati ancora maggiori, ossia i superammassi. Quando si leggono vecchi articoli tecnici sull'astronomia extragalattica si rimane colpiti dalle somiglianze fra le ipotesi di 50 anni or sono e i concetti meglio approfonditi di oggi. Quel che mancava ai nostri immediati predecessori erano strumenti di osservazione adeguati, come quelli che oggi forniscono i dati necessari per confermare alcune di quelle vecchie supposizioni. Benché le osservazioni nei raggi X, nell'ultravioletto, nell'infrarosso e nelle onde radio abbiano aperto nuove eccitanti finestre sull'universo, è giusto dire che le informazioni più importanti per la cosmologia sono state ottenute con telescopi che raccolgono luce nelle frequenze del visibile e dell'infrarosso vicino.

Già prima dell'invenzione del telescopio, gli osservatori erano in grado di vedere nel cielo notturno non solo pianeti e stelle, ma anche chiazze di luce dall'aspetto nebuloso. Con l'introduzione di telescopi sempre più potenti, nell'Ottocento alcune di tali nebulose furono risolte in singole stelle. Taluni astronomi sostennero che quelle nebulose dovevano essere giganteschi sistemi stellari indipendenti, esterni al nostro sistema. In ogni caso le posizioni di nebulose elencate nel *General Catalogue* (1864) di John Herschel e nel *New General Catalogue* (1888) di J. L. E. Dreyer presentano chiaramente dei raggruppamenti.

Di conseguenza gli astronomi che vedevano nelle nebulose dei sistemi esterni cominciarono a formulare ipotesi sulla loro tendenza a raggrupparsi in unità maggiori. Nel 1908 l'astronomo svedese C. V. L. Charlier delineò una cosmologia caratterizzata da una gerarchia di raggruppamenti. Identificò un certo numero di ammassi nebulari, e due fra i più grandi si trovavano nelle costellazioni Virgo e Coma Berenices. Nel 1922 l'astronomo britannico J. H. Reynolds osservò una fascia di nebulose che sembrava distendersi dall'Ursa Major alla Virgo passando per la Coma Berenices su un arco di ben 40 gradi nel cielo boreale. Reynolds riteneva che queste nebulose si trovassero all'interno del nostro sistema stellare, ma fu il primo a identificare quello che è riconosciuto oggi come il Superammasso locale, di cui fa parte anche la nostra galassia.

Alla metà degli anni venti, Edwin P. Hubble, del Mount Wilson Observatory, dimostrò che molte nebulose sono in realtà galassie esterne. Nel 1929 egli annunciò l'importante scoperta, fatta in collaborazione con Milton L. Humason, che quanto maggiore è la distanza di una galassia (giudicata sulla base della debolezza relativa della sua luce) tanto più la sua radiazione risulta spostata verso l'estremità rossa dello spettro elettromagnetico. Tali spostamenti verso il rosso indicano che le galassie si allontanano da noi (e fra loro) nel quadro di un'espansione generale dell'universo. Oggi il rapporto tra spostamento verso il rosso e distanza è detta legge di Hubble ed è alla base della moderna cosmologia d'osservazione.

Il valore dello spostamento verso il rosso, z, si ottiene sottraendo la lunghezza d'onda di riposo, o lunghezza d'onda non spostata verso il rosso, delle righe spettrali stellari di una galassia dalla lunghezza d'onda osservata e dividendo la differenza per la lunghezza d'onda di riposo. Il valore massimo dello spostamento verso il rosso trovato da Humason (verso la fine degli anni quaranta) fu di 0,2, equivalente a una velocità di recessione di 60 000 chilometri al secondo, ossia al 20 per cento della velocità della luce. Non c'è fra gli astronomi un accordo generale sul valore da attribuire al rapporto di Hubble, ma un valore plausibile fa corrispondere una

La regione centrale dell'ammasso di galassie nella costellazione di Coma Berenices, una piccola parte del grande superammasso comprendente non solo l'ammasso di Coma, ma anche un altro ammasso ricco, A1367, è mostrata qui in una rappresentazione a colori ricostruita da fotografie in bianco e nero. Nella fotografia originale si possono contare almeno 300 fra galassie ellittiche e galassie del tipo designato come S0, di luminosità superiore a un limite assegnato, e ciascuna di tali galassie è un gigantesco aggregato di decine di miliardi di stelle situate a una distanza di circa 300 milioni di anni luce da noi. In questa fotografia l'unico oggetto di spicco che non sia una galassia è una stella azzurra luminosa, leggermente al di sopra e a destra del centro, che è un oggetto vicino, appartenente alla nostra galassia. Le due gigantesche galassie ellittiche in prossimità del centro sono dominate dalla luce rossastra di stelle molto più vecchie. Questa immagine fu preparata al Kitt Peak National Observatory con l'aiuto di un monitor televisivo controllato da un calcolatore in cui si possono equilibrare con precisione colori appropriati a fotografie in bianco e nero eseguite separatamente nelle regioni rossa, verde e blu (o ultravioletta) dello spettro elettromagnetico. Le fotografie originali furono eseguite da uno degli autori (Thompson) con il telescopio Schmidt di 1,2 metri di Monte Palomar.

La galassia a spirale NGC 4535 si trova nella costellazione Virgo a una distanza dalla Terra approssimativamente di 60 milioni di anni luce. Essa fa parte del Superammasso locale di galassie, centrato sull'ammasso di Virgo. La nostra galassia si trova in un ammasso scarsamente popolato conosciuto come Gruppo locale, che oggi si ritiene appartenga al Superammasso locale. La tonalità rosso-arancione che si osserva nel nucleo di NGC 4535 indica un'abbondanza di stelle vecchie e fredde. I noduli azzurri nei bracci di spirale mettono in evidenza regioni in cui si sono formate di recente stelle giovani molto calde. Si ritiene che NGC 4535 assomigli alla nostra galassia per forma, dimensioni e luminosità. Questa immagine è una ricostruzione sulla base di fotografie eseguite da Thompson col telescopio Schmidt di Monte Palomar.

velocità di recessione di 75 chilometri al secondo con una distanza di un milione di parsec, ossia 3,26 milioni di anni luce. La galassia di Humason, col suo spostamento verso il rosso di 0,2, si trova perciò a 2,6 miliardi di anni luce di distanza. Gli oggetti più lontani oggi noti, i quasar, che hanno spostamenti verso il rosso anche di 3,5, si allontanano a quanto pare a velocità superiori al 90 per cento della velocità della luce. Si presume che si trovino a una distanza di circa 15 miliardi di anni luce.

Negli anni trenta sia Hubble sia Harlow Shapley, dello Harvard College Observatory, richiamarono l'attenzione sul fatto che, nell'emisfero galattico boreale (che comprende le costellazioni Virgo e Coma Berenices), le galassie luminose sono più numerose che non nell'emisfero australe. Hubble fotografò anche molte galassie così deboli da essere indotto a credere di aver trovato probabilmente un estremo del fenomeno del raggruppamento di galassie. «Per la prima volta - disse - la regione oggi osservabile con i telescopi esistenti potrebbe essere un campione rappresentativo dell'universo nella sua totalità.»

Un altro dei contributi di Hubble, che è pertinente alla nostra discussione, è il suo schema di classificazione delle forme di galassie. Hubble suddivise le galassie in due classi principali, ellittiche e a spirale, con varie categorie secondarie. Le galassie ellittiche variano fra i due estremi delle galassie sferiche e delle galassie lenticolari e sono generalmente prive di struttura interna. Le galassie a spirale, come la nostra, sono dischi appiattiti nei quali si distinguono di solito bracci di spirale (a meno che il disco non venga visto di taglio). In seguito fu identificato un terzo gruppo, quello delle galassie S0, che presentano proprietà intermedie.

Nel 1950 gli astronomi riuscirono a mettersi d'accordo sulle caratteristiche generali degli ammassi di galassie. Erano noti vari ammassi molto ricchi, il più grande dei quali era l'ammasso di Coma Berenices, comprendente più di mille galassie. Si è trovato che gli ammassi sono composti in prevalenza di galassie ellittiche e di galassie S0. All'interno degli ammassi si trovavano non più di metà di tutte le galassie; quanto alle altre, classificate come oggetti «di campo», si pensava fossero galassie isolate, per lo più a spirale, esterne agli ammassi. Alcuni astronomi suggerirono che la regione di Virgo poteva comprendere più di un singolo ammasso, ma il modello di Charlier, che ipotizzava una gerarchia di ammassi progressivamente maggiori, sembrava condannato dai conteggi, effettuati da Humason, di galassie deboli.

Gerard de Vaucouleurs, dell'Università del Texas ad Austin, che stava studiando sin dall'inizio degli anni cinquanta le galassie più luminose nell'emisfero galattico nord, fu il primo a definire e a descrivere il Superammasso locale. Il suo lavoro dimostra che esso è centrato sull'ammasso di Virgo, a una distanza da noi di circa 60 milioni di anni luce, e che possiede forse 50 ammassi esterni chiamati gruppi, oltre a galassie singole disseminate fra un gruppo e l'altro. La nostra galassia si trova in uno degli ammassi scarsamente popolati, chiamato Gruppo locale, cosicché è quasi certamente alla periferia del grande superammasso.

Un altro sviluppo importante nel corso degli ultimi trent'anni è la crescente consapevolezza del fatto che il Superammasso locale non è unico. Fra il 1950 e il 1954 fu eseguito un rilevamento fotografico dell'intero cielo boreale con il telescopio grandangolare Schmidt da 1,2 metri di Monte Palomar. Poco tempo dopo George O. Abell, dell'Università della California a Los Angeles, compilò un catalogo di 2712 ammassi ricchi di galassie. Abell sottolineò che molti ammassi sembravano membri di superammassi, con una media di cinque o sei ammassi ciascuno. Il suo suggerimento fu però contestato sulla base di un altro catalogo di ammassi compilato, sulla base del medesimo rilevamento fotografico, da Fritz Zwicky e colleghi al California Institute of Technology. Il catalogo di Zwicky faceva pensare che gli ammassi non fossero organizzati in ammassi di ordine superiore. Il disaccordo può essere risolto riconoscendo che gli ammassi quali sono definiti da Zwicky hanno generalmente dimensioni maggiori degli ammassi di Abell e comprendono vari centri di aggregazione di galassie.

Press'a poco nello stesso periodo, ma sulla base di un diverso rilevamento fotografico del cielo (compiuto al Lick Observatory), Jerzy Neyman, Elizabeth L. Scott e C. D. Shane dell'Università della California a Berkeley riferirono di avere individuato immense «nubi di galassie», termine da loro impiegato per definire i superammassi. Suggerirono anche provvisoriamente, sulla base del loro studio, che ogni galassia dell'universo appartenga a un ammasso e che non possano esistere galassie isolate. Più recentemente un'analisi completa e accurata di tutti i cataloghi di galassie disponibili è stata eseguita da P. J. E. Peebles e dai suoi colleghi alla Princeton University. Essi hanno quantificato l'intero spettro del raggruppamento di galassie e hanno concluso, fra l'altro, che gli ammassi tendono a raccogliersi l'uno vicino all'altro.

Un terzo importante sviluppo nell'analisi del fenomeno di raggruppamento di galassie è consistito, a partire dal 1950, nell'introduzione di rilevamenti su vasta scala degli spostamenti verso il rosso. Il primo passo in questa direzione è di misurare gli spostamenti verso il rosso di tutte le galassie la cui luminosità superi un dato limite inferiore in un'area prestabilita del cielo. Applicando agli spostamenti verso il rosso osservati la legge di Hubble, si può stimare la distanza di ciascuna galassia. Questa impostazione presenta sicuri vantaggi rispetto all'analisi statistica dei cataloghi di galassie esistenti, i quali forniscono solo le due coordinate angolari delle posizioni di galassie, ossia l'ascensione retta e la declinazione, cioè le due coordinate che possono essere misurate sul piano del cielo. Sulla base dei dati contenuti nei cataloghi esistenti, la terza dimensione, ossia la distanza, può essere dedotta solo in modo approssimato dalla debole luminosità della galassia. Nel metodo di rilevamento degli spostamenti verso il rosso, la distanza viene derivata esplicitamente dalla legge di Hubble. Lo svantaggio di questo metodo consiste nel fatto che, mentre una singola fotografia consente di stabilire migliaia di posizioni di galassie, gli spostamenti verso il rosso possono essere ottenuti solo uno per volta. I due metodi, nondimeno, sono complementari. Lo studio dei cataloghi consente di analizzare un gran numero di galassie in grandi volumi dell'universo; lo studio degli spostamenti verso il rosso fornisce particolari tridimensionali in volumi-campione molto più piccoli.

Rilevamenti su vasta scala di spostamenti verso il rosso sono stati resi possibili da importanti progressi nella strumentazione telescopica. Benché Hubble e Humason avessero accesso ai telescopi più potenti (il riflettore di 100 pollici di Mount Wilson e in seguito quello di 200 pollici di Monte Palomar), le emulsioni fotografiche disponibili erano lente rispetto a quelle attuali. I moderni spettrografi comprendono di solito un dispositivo elettronico che intensifica l'immagine almeno di un fattore 20, prima che essa entri nel rivelatore. Per certe osservazioni è possibile talvolta usare rivelatori numerici così sensibili da poter contare singoli fotoni. In conseguenza di questi e altri progressi, gli astronomi sono ora in grado di registrare in mezz'ora una quantità di informazioni spettrografiche pari a quella che Hubble e i suoi contemporanei potevano registrare solamente nell'arco di un'intera notte.

È di interesse storico il fatto che il primo rilevamento di spostamenti verso il rosso di galassie fu presentato a una conferenza tenuta nel 1960 sull'applicazione all'astronomia di tubi intensificatori di immagine. Con uno dei nuovi dispositivi, Nicholas U. Mayall, lavorando col riflettore da 120 pollici al Lick Observatory, aveva registrato lo spettro di 40 delle 82 galassie più luminose che si trovano a meno di quattro gradi dal centro dell'ammasso di galassie in Coma Berenices. Nel 1972 Herbert J. Rood e Thornton L. Page, della Wesleyan University, completarono ed estesero il rilevamento iniziale di Mayall. Altri spostamenti verso il rosso furono registrati da Eric C. Kintner, anch'egli della Wesleyan University, il quale analizzò poi il campione ampliato in collaborazione con Rood, Page e Ivan R. King dell'Università della California a Berkeley. Il loro articolo rappresenta il primo studio moderno, particolareggiato, degli spostamenti verso il rosso in un singolo ammasso di galassie. Essi riferirono che l'ammasso consta in prevalenza di galassie ellittiche e S0, alcune delle quali dotate di una velocità di più di 1000 chilometri al secondo, e che essi non erano in grado di fissare un limite alle dimensioni dell'ammasso.

Circa quattro anni dopo William G.

Tifft, dell'Università dell'Arizona, e uno di noi (Gregory) estesero il rilevamento di Coma Berenices sia a galassie più deboli sia a limiti angolari più vasti. Essi trovarono che l'ammasso in Coma termina a tre gradi dal centro, ma che un certo numero di galassie formano una proiezione simile a un braccio che si estende a ovest verso il vicino ammasso A1367, con cui sono forse connesse. (A1367 sta per ammasso N. 1367 nel catalogo di Abell. L'ammasso in Coma è A1656.) La nostra analisi sottolineò che i rilevamenti degli spostamenti verso il rosso forniscono non solo un quadro particolareggiato di ammassi di galassie lontani, ma anche informazioni importanti su galassie che possono trovarsi in primo piano. Poiché queste sembrano appartenere alle formazioni scarsamente popolate chiamate gruppi (o nubi, se sono ancor più rarefatte), un singolo rilevamento sistematico di spostamenti verso il rosso consente di identificare raggruppamenti di dimensioni diverse, dai più grandi ai più piccoli. In effetti, i campioni rarefatti in primo piano possono forse fornirci, sul modo in cui si formarono gli ammassi, informazioni non meno importanti degli ammassi ricchi più vistosi. La nostra analisi richiamò l'attenzione anche sulla scarsità delle galassie di campo.

In un settore di ricerca in rapido sviluppo, osservazioni uguali o simili vengono fatte spesso indipendentemente da vari scienziati. Accadde così che Rood e Guido L. Chincarini, dell'Università dell'Oklahoma, che stavano studiando galassie a ovest dell'ammasso in Coma Berenices, trovassero che il braccio proteso verso ovest dell'ammasso in Coma era ancora rilevabile a una distanza di più di 14 gradi. Essi suggerirono inoltre che l'ammasso in Coma e A1367 fossero uniti da un ponte di galassie. A questo punto noi due (Gregory e Thompson) iniziammo un rilevamento sistematico che estese la rassegna di Coma, verso ovest, sino all'ammasso A1367. La nostra analisi più completa confermò l'esistenza di una connessione, una sorta di ponte, fra i due ammassi. Uno studio che aveva avuto origine con l'osservazione, per opera di Mayall, di 40 galassie in un'area di 16 gradi quadrati, fu estesa così a 238 galassie in un'area di 260 gradi quadrati e nel corso di questo processo fu accertata l'esistenza di un vero superammasso.

L'ammasso in Coma si trova in prossimità del polo esteso della nostra galassia, a una novantina di gradi dal velo di polvere e di gas che limita la visibilità nel piano centrale della Galassia. Nel nostro studio del superammasso Coma-A1367, decidemmo di procurarci gli spettri di tutte le galassie di luminosità maggiore della magnitudine 15, una magnitudine circa un milione di volte più debole di quella di Vega, la stella più luminosa nel cielo boreale. Quando le galassie del nostro campione vengono rappresentate su una carta bidimensionale, come appaiono in cielo, si osservano due concentrazioni principali: l'ammasso in Coma, nell'angolo nordorientale, e l'A1367 nell'angolo sudoccidentale *(si veda l'illustrazione a pagina 106)*. A parte ciò si ha la forte impressione visiva che la carta sia costituita da molte galassie isolate, distribuite in modo più o meno casuale fra i due centri.

Se ora riportiamo in un grafico i risultati del rilevamento degli spostamenti verso il rosso in modo da illustrare come le stesse galassie siano distribuite nella terza dimensione, ossia secondo la distanza, emerge un quadro del tutto diverso. A questo scopo è sufficiente usare due sole coordinate di posizione: la distanza radiale (derivata dagli spostamenti verso il rosso) e gli angoli in direzione est-ovest in cielo *(si veda l'illustrazione a pagina 108)*. Con questa rappresentazione le galassie ci appaiono distribuite in modo molto meno uniforme.

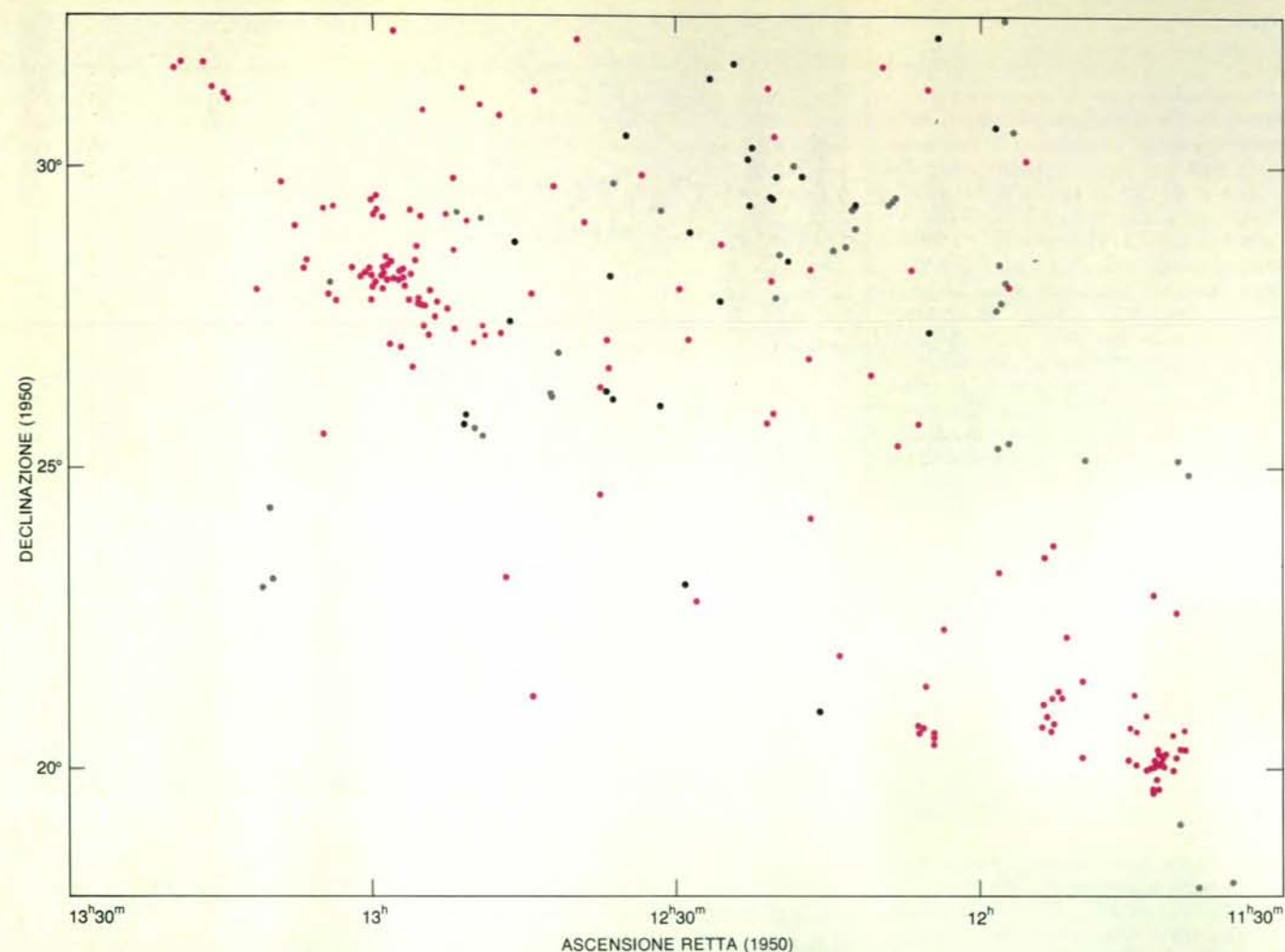
Vicino alla Galassia si trovano alcuni raggruppamenti minori, rappresentati dal vertice del cuneo. L'elemento più imponente è però una regione densamente popolata che si distende per l'intera estensione della carta, a una distanza di 315 milioni di anni luce da noi. Tale raggruppamento può essere definito come un superammasso, comprendendo i due ammassi ricchi, quello in Coma e A1367, e vari ammassi meno popolosi che formano congiuntamente una struttura megagalattica continua, estesa da un capo all'altro per più di 70 milioni di anni luce.

Sorprendentemente, la carta presenta anche alcuni «vuoti», regioni apparentemente prive di galassie. All'epoca in cui terminammo lo studio eravamo sì convinti che i vuoti fossero reali, ma avevamo dubbi circa la loro universalità, e pensavamo che potessero essere una peculiarità di questa regione del cielo.

Poiché i superammassi possono essere altrettanto diversi fra loro per struttura e contenuto quanto le singole galassie, vorremmo conoscere altri esempi, oltre al superammasso in Coma e ad A1367, prima di fare delle generalizzazioni sulle proprietà dei superammassi. Almeno altri tre grandi sistemi sono attualmente allo studio, e tutti presentano caratteristiche diverse di estremo interesse. Recentemente uno di noi (Thompson), in collaborazione con Chincarini, Rood, Tifft e Massimo Tarenghi, lavorando con i telescopi da due metri dello Stewart Observatory dell'Università dell'Arizona e del Kitt Peak National Observatory, ha studiato la regione attorno all'ammasso di Hercules. Ancora una volta i dati d'osservazione dimostrano la presenza di un superammasso complesso che occupa una larga banda, a una distanza compresa fra 400 e 600 milioni di anni luce dalla Terra. A differenza del superammasso Coma-A1367, il sistema di Hercules non è dominato da uno o due ammassi. Ciononostante assomiglia al superammasso Coma-A1367 in quanto in primo piano presenta una vasta regione vuota. L'aspetto forse più sorprendente del sistema di Hercules, però, è che i suoi ammassi più densi sono dominati da galassie a spirale anziché da galassie ellittiche. Questa peculiarità da sola è già sufficiente a renderlo degno di nota.

Il terzo superammasso allo studio si trova nella regione del cielo che coincide con le costellazioni Perseus e Pisces. Esso ha l'aspetto di un lungo filamento esteso per oltre 40 gradi in cielo, dal ben noto ammasso di galassie in Perseus a un piccolo gruppo di galassie in prossimità della galassia ellittica NGC 383. Le nostre nuove osservazioni, compiute in collaborazione con Tifft, indicano che la profondità di quello che appare come un filamento non è maggiore della sua larghezza. Possiamo presumere di conseguenza non solo che il sistema sia un vero filamento, ma anche che le singole galassie all'interno di tale struttura abbiano moti casuali di velocità piuttosto bassa. Ci sono anche indizi del fatto che molte delle singole galassie nel sistema Perseus-Pisces abbiano piani di rotazione o paralleli all'asse del filamento del superammasso o perpendicolari ad esso. Quest'osservazione, se confermata, potrebbe dirci qualcosa sul meccanismo di formazione di galassie e superammassi.

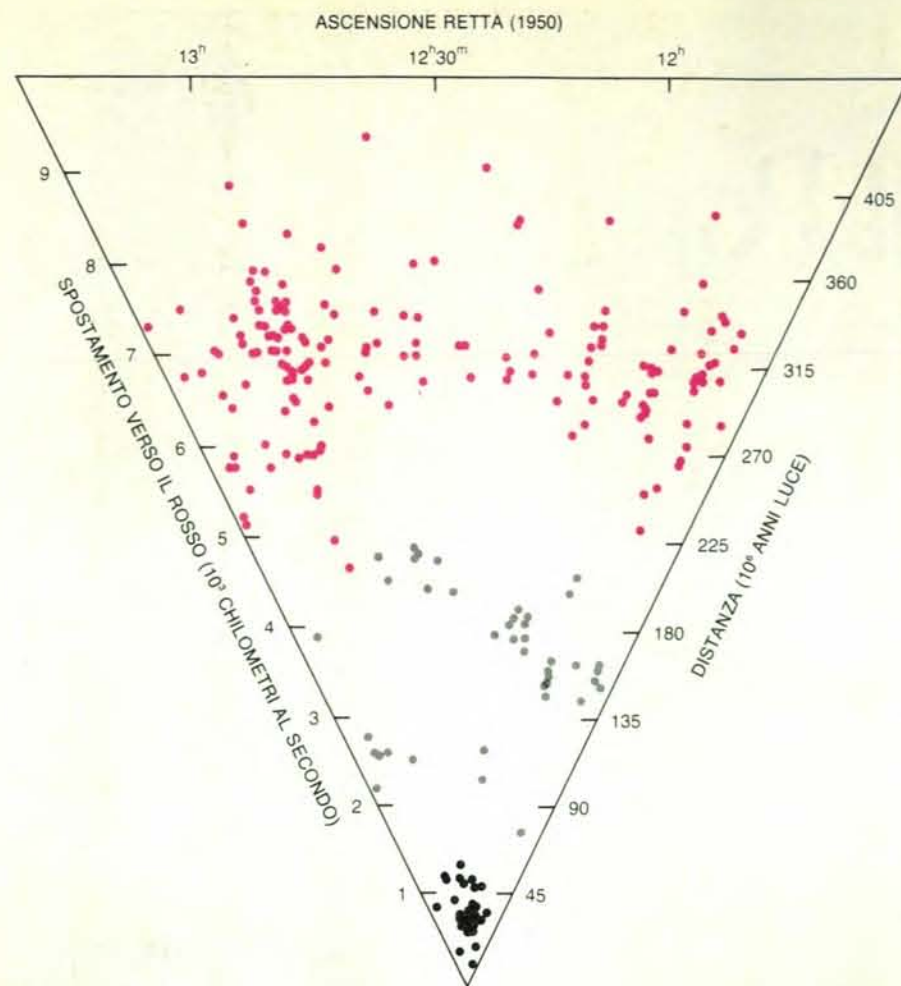
I tre rilevamenti sistematici di spostamenti verso il rosso di galassie coprono solo il 2 per cento circa del cielo. Gruppi di studio in vari osservatori stanno cercando di ottenere una visione più ampia e organica del fenomeno dei superammassi. Per esempio, Jaan Einasto, Mikhel Jôeveer, Enn Saar e S. Tago, in Estonia, che hanno scoperto indipendentemente il superammasso di Perseus e l'esistenza di vuoti, hanno analizzato il più esteso catalogo esistente di spostamenti verso il rosso di galassie. Nonostante il loro catalogo non abbia l'abbondanza di particolari dei nostri recenti rilevamenti sistematici, ha consentito loro di verificare su scala maggiore le medesime caratteristiche riscontrate con i me-

Su questa proiezione omolografica della sfera celeste è indicata la posizione di quattro superammassi. L'ammasso Hydra-Centaurus, nell'emisfero australe, non è stato ancora studiato in misura approfondita. Il primo superammasso identificato fu il sistema Coma Berenices-A1367. Gli autori stimano che esso possa contenere assai più di un milione di galassie: di 280 di queste galassie è stato a tutt'oggi misurato lo spostamento verso il rosso. Questo superammasso occupa un volume di almeno 10^{23} anni luce cubi. I tre quadratini neri identificano regioni in cui un rilevamento recente ha rivelato un'assenza pressoché totale di galassie con velocità di recessione comprese fra 12 000 e 18 000 chilometri al secondo, che corrispondono all'intervallo fra 520 e 780 milioni di anni luce. Se la regione compresa fra quei tre siti di campioni dovesse risultare anch'essa mancante di galassie, corrisponderebbe a un vuoto di circa 30 × 10^{24} anni luce cubi.

La distribuzione delle galassie nella regione del cielo comprendente gli ammassi di Coma Berenices e A1367 è stata rappresentata cartograficamente dagli autori nel corso del loro vasto rilevamento sistematico di spostamenti verso il rosso. Le galassie più lontane sulla carta sono segnate in colore, le più vicine sono in nero mentre quelle a distanze intermedie sono rappresentate in grigio. La densa concentrazione di galassie a un'ascensione retta di 13 ore e a una declinazione di 28 gradi rappresenta la regione centrale dell'ammasso in Coma. L'aggregazione minore corrispondente all'ascensione retta di 11 ore e 40 minuti e alla declinazione di 20 gradi rappresenta l'ammasso A1367. Poiché le galassie altrove in cielo sembrano disseminate in modo abbastanza uniforme, gli astronomi furono indotti a pensare che la maggior parte delle galassie fossero disseminate nello spazio in modo casuale. Una veduta diversa della medesima area del cielo viene rappresentata a pagina 108.

todi di rilevamento più particolareggiati.

Analogamente, Chincarini e Rood hanno analizzato la distribuzione di galassie spirali giganti osservata per la prima volta da Vera C. Rubin, da W. Kent Ford, Jr., e dai loro colleghi del Dipartimento di geomagnetismo della Carnegie Institution di Washington. Il rilevamento Rubin-Ford copre l'intero cielo ma presenta solo pochi particolari in ogni regione. Ciononostante, Chincarini e Rood hanno potuto verificare l'esistenza dei tre superammassi da noi descritti e di identificare un'altra nuova struttura nel cielo australe: il superammasso Hydra-Centaurus. Le ricerche di Chincarini e Rood e di Einasto, Jôeveer, Saar e Tago suggeriscono in modo convincente che i superammassi si estendano molto oltre le regioni coperte dai nostri rilevamenti di spostamenti verso il rosso. Secondo i loro calcoli il superammasso Coma-A1367 e quello di Perseus possono occupare un'area di cielo dieci volte maggiore di quella da noi prudentemente proposta.

Quest'ipotesi ha ricevuto ulteriore sostegno da un rilevamento eseguito da Robert P. Kirshner dell'Università del Michigan, da Augustus Oemler, Jr., della Yale University, da Paul L. Schechter del Kitt Peak National Observatory e da Stephen A. Shectman degli osservatori di Mount Wilson e Las Campañas. Questo rilevamento copre tre piccole regioni dell'emisfero galattico nord. In ciascuna regione essi hanno trovato galassie con spostamenti verso il rosso simili a quelli osservati nel superammasso Coma-A1367, e ritengono anche di avere scoperto un vuoto immenso, il cui volume potrebbe essere prossimo a 30×10^{24} anni luce cubi. Nelle loro piccole aree di ricerca, centrate in prossimità del polo galattico nord, pare siano quasi completamente assenti galassie con spostamenti verso il rosso compresi fra 12 000 e 18 000 chilometri al secondo. In circa quattro gradi quadrati di cielo, dove si aspettavano di trovare circa 25 galassie con spostamenti verso il rosso in quell'ordine di valori, ne hanno trovato solo una. Le dimensioni di quel vuoto, che sarebbe di gran lunga il più grande finora noto, sono state calcolate supponendo che la regione compresa nello spazio delimitato dai punti scelti per il campionamento sia ugualmente vuota di galassie. Il vuoto si trova a una distanza dalla Terra compresa fra 520 e 780 milioni di anni luce.

Sulla base dei rilevamenti eseguiti a tutt'oggi abbiamo riportato in grafico la distribuzione tridimensionale delle galassie nei tre superammassi ben definiti: Coma-A1367, Hercules e Perseus *(si veda l'illustrazione a pagina 110)*. In questa rappresentazione la nostra galassia si trova al centro. I tre rilevamenti ci dischiudono finestre a forma di cuneo attraverso le quali possiamo gettare uno sguardo nella vastità del cosmo. La tendenza delle galassie a raccogliersi in unità maggiori, ci appare onnipresente. Non si può inoltre più dubitare dell'esistenza di vuoti, sulla quale inizialmente eravamo scettici. L'universo avrebbe potuto orga-

Una «visione in piano» del superammasso Coma Berenices-A1367 può essere disegnata dal rilevamento degli spostamenti verso il rosso compiuto dagli autori. Le galassie sono le stesse rappresentate a pagina 106, ma qui sono distribuite secondo le loro velocità di recessione, quali sono desunte dagli spostamenti verso il rosso. La costante di Hubble, che stabilisce una relazione fra velocità e distanza, fornisce la scala che compare sul lato destro del triangolo. Il fattore di conversione adottato fa corrispondere a una velocità di recessione di 1000 chilometri al secondo una distanza di 43,5 milioni di anni luce (ossia 75 chilometri al secondo per milione di parsec). La nostra galassia dovrebbe trovarsi sul vertice inferiore del triangolo, a una distanza di zero anni luce. Si può vedere ora che l'ammasso in Coma (che si trova a un'ascensione retta di 13 ore, con una velocità di recessione di circa 7000 chilometri al secondo) e l'ammasso A1367 (che è a 11 ore 40 minuti e ha la medesima velocità di recessione) sono semplicemente gli estremi riccamente popolati di una banda continua di galassie che si distende in cielo a una distanza di circa 315 milioni di anni luce dalla nostra galassia. Questa banda di galassie è il superammasso Coma-A1367. Si osservino anche gli enormi vuoti là dove il rilevamento non ha riscontrato la presenza di galassie.

nizzarsi in modo che gli spazi fra gli ammassi fossero non totalmente vuoti, ma riempiti da piccoli gruppi di galassie. I vuoti sono invece evidentemente parte integrante del processo che ha condotto alla formazione di raggruppamenti e superammassi.

Lo studio dei superammassi non è un dominio esclusivo dell'astronomia ottica; anche la radioastronomia e l'astronomia a raggi X stanno apportando contributi fondamentali. I radioastronomi sono in grado di rivelare la presenza di gas intergalattico dimostrando che alcune radiosorgenti in ammassi e superammassi sono state distorte da quello che è presumibilmente un gas di bassa densità e alta temperatura. Se il gas riempie interi superammassi nello stesso modo in cui riempie le regioni più dense degli ammassi, il suo contributo alla massa totale del superammasso potrebbe essere enorme.

L'astronomia a raggi X ha già permesso di scoprire in direzione di superammassi remoti gas estremamente caldo. Non è chiaro però se l'emissione provenga esclusivamente dai nuclei di ammassi luminosi o se una parte di essa abbia origine nelle regioni comprese fra i nuclei di ammassi. Jack O. Burns, Jr., dell'Università del New Mexico e uno di noi (Gregory) stanno combinando i dati sugli spostamenti verso il rosso ottenuti al Kitt Peak National Observatory, le carte radio di sorgenti distorte realizzate con i dati forniti dal radiotelescopio Very Large Array a Socorro, nel New Mexico, e le immagini a raggi X ottenute dal satellite Einstein, per esaminare un campione completo di ammassi.

Altri astronomi stanno applicando i metodi della radioastronomia per eseguire i loro rilevamenti degli spostamenti verso il rosso. Questi vengono determinati osservando lo spostamento nella riga di radioemissione a 21 centimetri dell'idrogeno neutro (non ionizzato) nello spazio interstellare. Un rilevamento sistematico, condotto da J. Richard Fisher e R. Brent Tully, dell'Università di Hawaii a Manoa, ha registrato cartograficamente le galassie del Superammasso locale sino a raggiungere, verso l'esterno, il grande vuoto più prossimo. Il radiotelescopio più sensibile di cui si disponga per questi studi è l'antenna da 303 metri ad Arecibo, nel Portorico, dove si sta lavorando su tutt'e tre i superammassi da noi descritti. Fra gli astronomi che lavorano a questi progetti vi sono Chincarini, Thomas M. Bania, Riccardo Giovanelli, Martha P. Haynes, Trinh Xuan Thuan e uno di noi (Thompson). Queste osservazioni sono preziose perché forniscono non solo valori dello spostamento verso il rosso per altre galassie, ma anche dati sull'idrogeno neutro in varie posizioni all'interno dei superammassi. Benché questi studi non siano abbastanza avanzati da fornire nuove conclusioni sulle condizioni vigenti all'interno dei superammassi, essi promettono molto per il futuro.

È diventato oggi estremamente chiaro, dai rilevamenti sistematici degli spostamenti verso il rosso, che l'attuale distribuzione delle galassie è estremamente eterogenea sino a una distanza di varie centinaia di milioni di anni luce. Pare probabile che questa disomogeneità si estenda sino a miliardi di anni luce e caratterizzi l'intero universo. Dobbiamo aggiungere però che l'universo potrebbe contenere molta materia non luminosa. Molti oggi stanno ripetendo e formulando ipotesi in merito alla possibile esistenza di questa materia e al suo possibile volume.

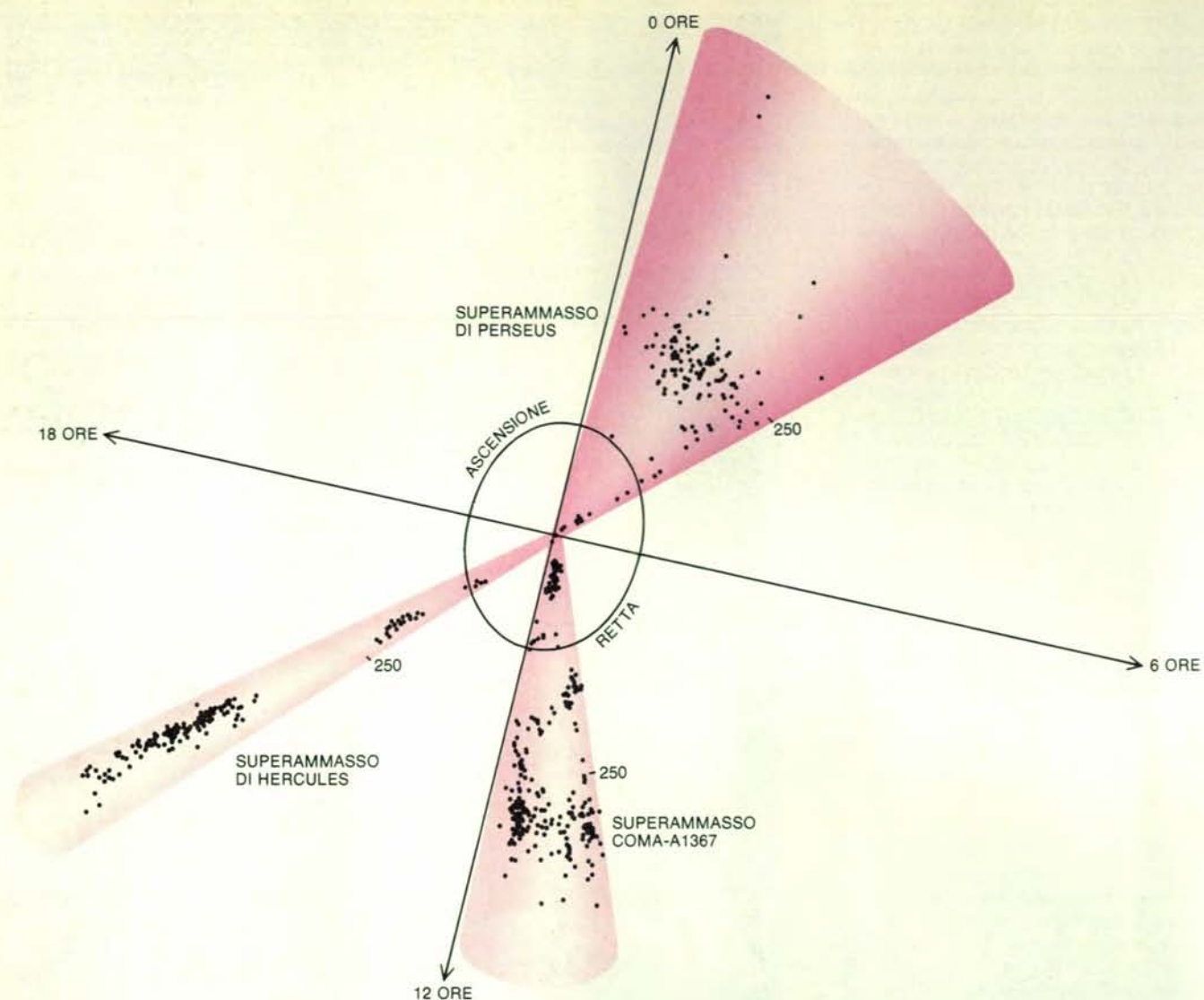
Oggi l'universo è caratterizzato da una mancanza di omogeneità, ma vi sono indizi del fatto che in epoche molto remote fosse omogeneo. Questi indizi consistono nel fatto che la debole radiazione di fondo in cui la Terra è immersa a lunghezze d'onda radio nella gamma delle microonde è notevolmente uniforme per quanto concerne la direzione. Secondo l'opinione più diffusa la radiazione di fondo rappresenterebbe il residuo, espanso e raffreddato, del primordiale universo caldo. La radiazione di fondo a microonde è dappertutto omogenea (le eterogeneità non superano una parte su 3000). Recentemente, tuttavia, sono state rilevate alcune disomogeneità di piccola ampiezza, che si estendono però su vasti angoli di cielo.

È possibile ricostruire il percorso evolutivo dall'omogeneità iniziale al ricco assortimento delle strutture attuali? Noi crediamo che l'assenza di galassie isolate e la presenza di grandi vuoti possano fornire indizi diretti per stabilire i tempi relativi di formazione di galassie, ammassi di galassie e ammassi di ammassi. Ci sono due ipotesi in concorrenza fra loro. Il modello più convenzionale suppone che le singole galassie siano sorte da un «brodo» primordiale quasi omogeneo. La difficoltà principale in cui questo modello incorre è che non riesce a spiegare in modo soddisfacente come l'universo sia passato dal suo stato omogeneo iniziale allo stato in cui la materia si raccolse in galassie. Il modello suppone che, una volta formatesi le galassie, piccole irregolarità nella loro distribuzione sarebbero state lentamente amplificate dall'azione delle forze gravitazionali a lungo raggio. Il risultato finale di tali amplificazioni sarebbero i superammassi visibili oggi.

Una spiegazione teorica alternativa fu suggerita nel 1972 da due astronomi sovietici, Jakov Zel'dovič e Rašid Sunjaev. Nel loro modello il gas dell'universo primordiale non si condensò immediatamente in stelle e galassie. Irregolarità lievi, ma su larga scala, nella distribuzione generale del gas si intensificarono in conseguenza dell'attrazione gravitazionale e divennero sempre più irregolari. Infine il gas divenne abbastanza denso da raccogliersi in vasti strati di materiale (ai quali fu appioppato il nomignolo di *pancakes*, frittelle), che poi si frammentarono in galassie. Secondo quest'ipotesi, ammassi e superammassi si formano dapprima come concentrazioni di gas, e solo in un secondo tempo appaiono le galassie.

L'uno o l'altro di questi modelli trovano conferma nelle osservazioni da noi compiute di superammassi? Poiché il modello di Zel'dovič e Sunjaev richiede che tutte le galassie si siano formate in ammassi o in superammassi, le galassie di campo, o vagabonde casuali, dovrebbero essere rare. Se è giusto invece il modello convenzionale e le galassie possono avere origine quasi dappertutto a caso, per raccogliersi solo in seguito in conseguenza della gravità in gruppi o ammassi, le galassie isolate dovrebbero essere piuttosto comuni. Di fatto le uniche popolazioni di galassie isolate, da noi scoperte nei nostri rilevamenti sistematici di spostamenti verso il rosso, sono galassie disseminate all'interno dei confini di superammassi. I vuoti, inoltre, sono veramente vuoti.

Noi suggeriamo che le galassie isolate, disseminate all'interno di superammassi, fossero un tempo membri di piccoli gruppi successivamente disgregati da collisioni all'interno dei superammassi densi. Pare realistico supporre che in una certa epoca tutte le galassie fossero membri di gruppi o di ammassi. In definitiva, la distribuzione osservata di galassie all'interno di superammassi e l'esistenza di grandi vuoti fra superammassi sono del tutto in accordo col modello di Zel'dovič e Sunjaev. I fautori del modello alternativo ricercano una conferma in simulazioni al calcolatore che, nelle loro speranze, dovrebbero dimostrare come irregolarità su piccola scala possano accentuarsi sempre più in conseguenza di processi casuali. Non è chiaro come si risolverà questa disputa.

Descrivendo il superammasso in forma di filamento di Perseus-Pisces suggerimmo la possibilità che gli assi di rotazione di alcune galassie fossero correlati non

Tre direzioni in cielo sono state studiate a tutt'oggi, in rilevamenti sistematici di spostamenti verso il rosso di galassie. In questa proiezione il sistema solare si trova all'apice di tre coni comprendenti le galassie studiate in ciascuna delle tre regioni con superammassi di galassie. Il numero adiacente a ciascun cono dà la distanza dal sistema solare in milioni di anni luce. La scala tiene conto degli effetti di scorcio nella proiezione. Finora solo il 2 per cento circa dell'intero cielo è stato sottoposto a rilevamenti particolareggiati di questo tipo delle galassie.

solo con gli assi di rotazione di altre galassie, ma forse anche con la struttura generale del filamento del superammasso. L'idea ha ricevuto sostegno da uno studio recente di Mark T. Adams, Stephen E. Strom e Karen M. Strom, del Kitt Peak National Observatory, i quali hanno trovato correlazioni di rotazione simili nei dati combinati relativi a vari ammassi appiattiti. Se tali correlazioni trovassero conferma, i fautori del modello convenzionale di formazione delle galassie probabilmente dovrebbero fronteggiare ostacoli insormontabili per riuscire a spiegarle. I processi statistici casuali del modello convenzionale non possono condurre a un moto di rotazione organizzato su vasta scala. Tali correlazioni invece verrebbero facilmente spiegate nel modello di Zel'dovič e Sunjaev.

Quali sono le prospettive di risolvere tali problemi nel prossimo futuro? Una fra le vie di indagine più promettenti è quella del continuo miglioramento nelle tecniche per misurare la radiazione di fondo primordiale a microonde. Le sia pur lievissime irregolarità osservate in questa radiazione attestano l'esistenza di una struttura in quest'universo già in epoca molto vicina all'origine. I limiti superiori dell'omogeneità osservata sono vicini a quelli richiesti per poter mettere a confronto i due modelli alternativi di formazione delle galassie. Forse nei prossimi dieci anni ci renderemo conto che nessuno dei due modelli è soddisfacente, nel qual caso gli astrofisici dovranno rimeditare completamente il problema.

I nostri commenti finali riguardano il concetto stesso di superammasso di galassie. Innanzitutto, superammasso è un termine adeguato? Secondo molti miei colleghi il termine dovrebbe essere riservato ad aggregati tenuti assieme da legami gravitazionali. Dalle osservazioni non è chiaro se questa condizione sia soddisfatta. La nostra opinione è che il termine superammasso descriva in modo appropriato quello che oggi è un aggregato di galassie ben separato da ogni altra struttura simile, senza che ciò debba implicare necessariamente l'esistenza di relazioni dinamiche fra membri di superammassi.

Una seconda osservazione concerne l'universalità dei superammassi. Oggi è accertato che ogni ammasso vicino riccamente popolato compreso nel catalogo di Abell fa parte di un superammasso. Ipotizziamo perciò che una condizione necessaria per la formazione di un ammasso ricco sia la presenza di un ammasso compagno. Infine vogliamo che il lettore si faccia un'idea appropriata delle dimensioni dei superammassi. Il superammasso Coma-A1367 dista più di 300 milioni di anni luce dalla Galassia, eppure, anche osservato da una distanza così enorme, ha ancora un'estensione di almeno 20 gradi in cielo, nelle costellazioni di Coma Berenices e Leo. Chincarini e Rood sostengono, però, che lo stesso superammasso potrebbe essere anche 10 volte più grande. Strutture organizzate di questa vastità aprono ampi orizzonti di studio ad astronomi e cosmologi.

Il serpente a sonagli degli Stati Uniti sudoccidentali può localizzare deboli sorgenti di radiazioni infrarosse per mezzo degli organi a fossetta, una coppia di cavità situate frontalmente e in posizione abbassata rispetto agli occhi. Nei serpenti a sonagli e in altri viperidi, la radiazione viene rivelata da una sottile membrana sensibile al calore sottesa nella parte posteriore di ogni fossetta. Il campo visivo si estende da 25 gradi oltre l'asse cefalocaudale dell'animale a 105 gradi di lato e da 45 gradi al di sotto a 60 gradi al di sopra dell'asse trasversale.

Il pitone reticolato, originario dell'Asia meridionale, ha 13 coppie di fossette nelle squame che stanno sopra e sotto la bocca. In questa fotografia si possono osservare cinque fossette a sinistra della narice e una a destra. Lungo il labbro inferiore, al di sotto e indietro rispetto all'occhio, si trovano sei altre fossette. Un'analoga disposizione si può ritrovare in altri serpenti appartenenti alla famiglia dei boidi, compresi i boa arboricoli. La superficie posteriore di ogni fossetta è estremamente sensibile all'energia termica. Ogni fossetta ha un campo di sensibilità leggermente differente dalle altre. Altri serpenti della stessa famiglia, compreso il *Boa constrictor*, invece di vere e proprie fossette, presentano sul muso delle squame sensibili all'infrarosso. Entrambe le fotografie della pagina sono state scattate da uno degli autori (Newman).

La visione infrarossa nei serpenti

Alcuni serpenti possono localizzare le radiazioni infrarosse che, integrate nel cervello con le informazioni relative al visibile, sono in grado di fornire loro un quadro unico, ad ampio spettro del mondo circostante

di Eric A. Newman e Peter H. Hartline

I serpenti a sonagli possono colpire con precisione anche di notte, quando le loro prede sono nascoste dal buio. Il colpo è guidato dal calore, cioè dalla radiazione infrarossa emessa dai bersagli a sangue caldo. La radiazione viene percepita da uno straordinario sistema che permette al serpente a sonagli di «vedere» in una regione dello spettro elettromagnetico in cui gli animali generano energia radiante.

Gli «occhi» per l'infrarosso del serpente a sonagli sono gli organi a fossetta, una coppia di profonde cavità che si aprono sui lati del capo, al di sotto e poco più avanti degli occhi. Gli organi a fossetta sono forniti di numerose fibre nervose sensibili al calore collegate con il cervello. Tutti i serpenti della sottofamiglia crotalini hanno organi a fossetta e sono sensibili alle radiazioni infrarosse. In America Settentrionale questi serpenti comprendono, oltre al serpente a sonagli, il moccassino acquatico e il testa di rame. Anche i pitoni e molti altri membri della famiglia boidi (i serpenti boa), lontanamente imparentati con i viperidi, hanno delle fossette sensibili al calore. Contrariamente ai viperidi, però, questi hanno numerose fossette (fino a 13 paia) sulle squame intorno alla bocca.

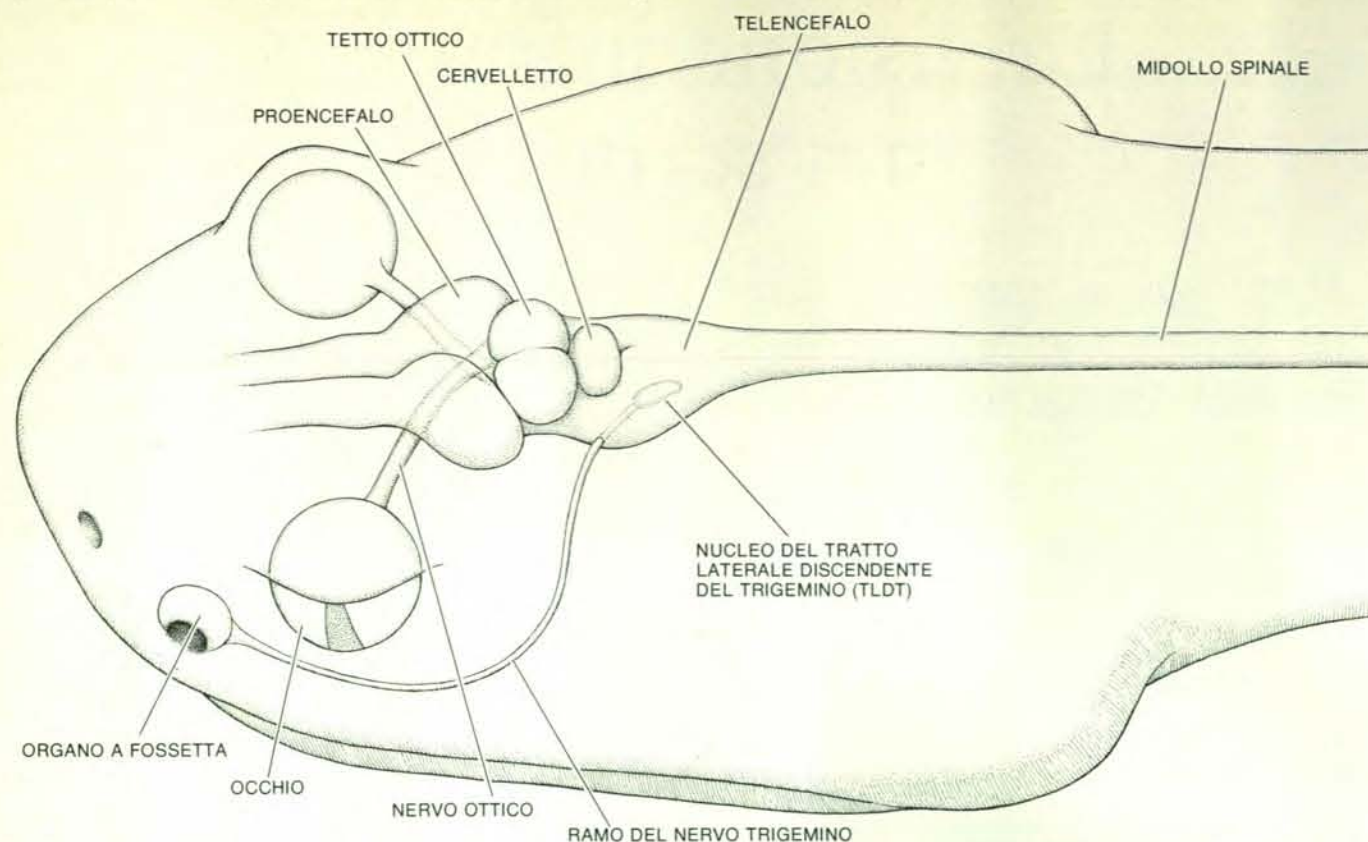
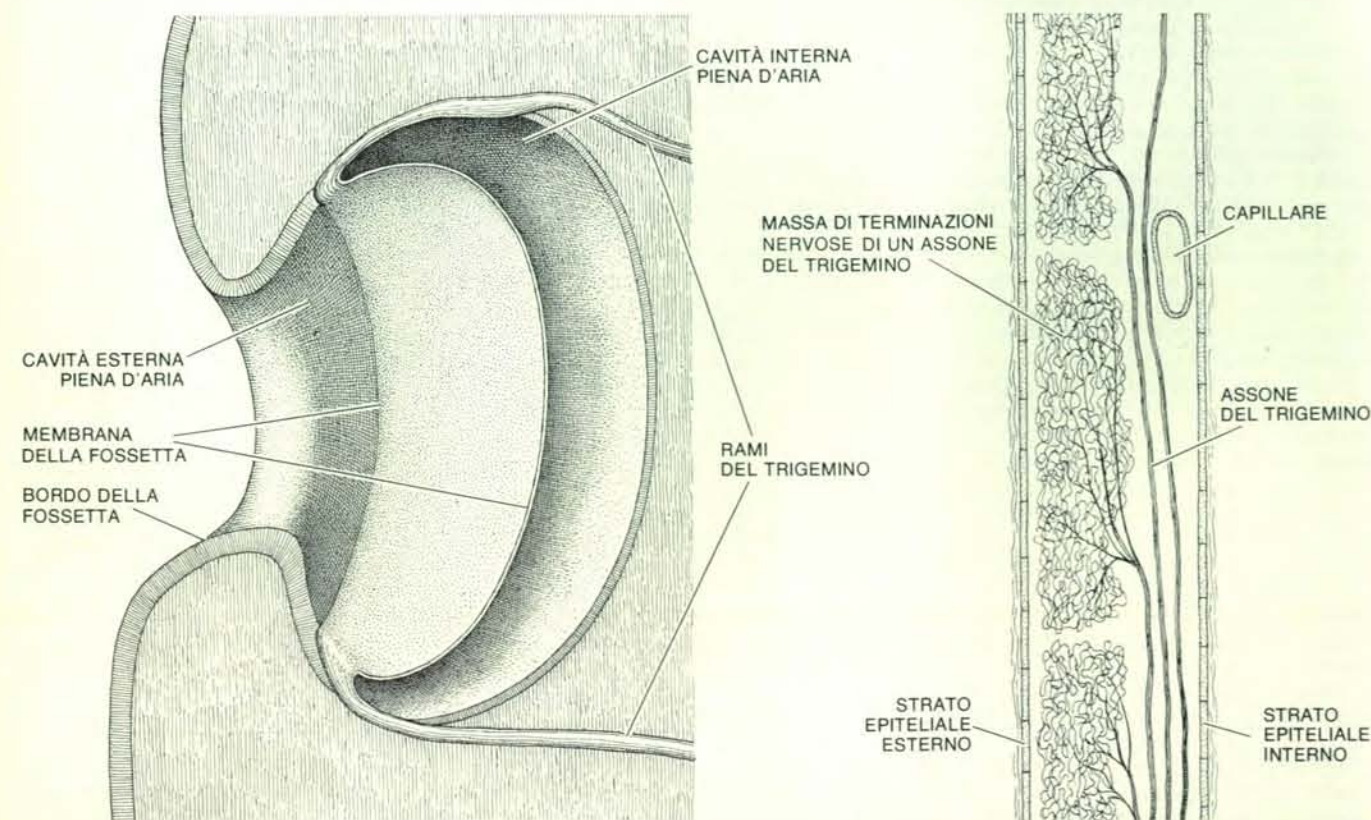
La sensibilità all'infrarosso dei viperidi e dei boidi sembra che si sia evoluta dal sistema sensoriale somatico. Nell'uomo e negli altri mammiferi questo sistema interessa, tra le altre, le sensazioni di tatto, pressione, temperatura e dolore. Nei mammiferi è il nervo trigemino a condurre i messaggi sensoriali dall'area facciale al cervello. Nei serpenti sensibili all'infrarosso una notevole parte del trigemino è destinata agli organi a fossetta e alla sensibilità all'infrarosso. Nel corso dell'evoluzione il sistema sensoriale somatico dei viperidi e dei boidi si è sviluppato in un senso molto simile alla vista. I viperidi e i boidi hanno anche sviluppato nel cervello dei nuclei, o ammassi di corpi cellulari nervosi, specializzati per ricevere ed elaborare le informazioni sensoriali provenienti dagli organi a fossetta. I nuclei trasmettono le informazioni al tetto ottico del mesencefalo, una struttura nota per il ruolo che ha nella visione e nella rappresentazione spaziale dell'informazione sensoriale.

Gli esseri umani imparano come gli altri animali ad associare varie modalità sensoriali con oggetti reali di vario tipo. Ad esempio, sia la vista di un uccello sia il suono del suo canto aiutano a localizzarlo e a identificarlo. Analogamente, il tetto ottico dei serpenti sensibili all'infrarosso combina i segnali infrarossi provenienti dagli organi a fossetta con i segnali visivi provenienti dagli occhi. L'integrazione delle informazioni visive e infrarosse nel tetto fornisce ai viperidi e ai boidi una visione unica del mondo, in cui le immagini visive e infrarosse vengono paragonate e contrapposte.

Sebbene gli organi a fossetta sensibili all'infrarosso fossero stati descritti da molto tempo dagli anatomisti, la loro funzione era rimasta oscura fino agli anni trenta, quando G. Kingsley Noble e A. Schmidt dell'American Museum of Natural History di New York dimostrarono che le fossette percepiscono il calore. Essi scoprirono che i serpenti a sonagli sono in grado di distinguere tra una lampadina calda (coperta con un panno opaco) e una fredda. I serpenti si orientavano e si lanciavano verso le lampadine calde finché le loro fossette rimanevano scoperte, mentre se queste venivano bloccate ignoravano tutte le lampadine, sia calde sia fredde.

Negli anni cinquanta Theodore H. Bullock e i suoi colleghi dell'Università della California a Los Angeles hanno dimostrato che gli organi a fossetta funzionano percependo le radiazioni infrarosse emesse dagli oggetti caldi. Questi studiosi hanno controllato gli impulsi nervosi negli assoni (lunghe fibre di cellule nervose) del trigemino collegati con le fossette e hanno trovato che gli impulsi vengono attivati da un oggetto caldo presentato al serpente tanto in piena luce quanto in una stanza buia. Gli assoni non rispondono, anche in una stanza illuminata, se la fossetta viene esposta allo stesso oggetto una volta che questo si è raffreddato. Inoltre, l'attività dell'assone in risposta a un oggetto caldo scompare se tra la fossetta e l'oggetto si inserisce un filtro di vetro che assorbe il calore, anche se il vetro lascia passare quasi tutta la luce visibile. Nell'esperimento complementare, in cui veniva inserito un filtro che non lasciava passare le radiazioni visibili, ma permetteva il passaggio della maggior parte delle radiazioni infrarosse, la risposta risultava solo debolmente attenuata. Si è così dimostrato in modo decisivo che l'organo a fossetta è sensibile alle radiazioni di lunghezze d'onda nell'infrarosso.

Ma quanto è sviluppata la sensibilità all'infrarosso nel serpente a sonagli? Con quale precisione un serpente a sonagli è in grado di localizzare un oggetto caldo? Qual è la sensibilità dei rivelatori di infrarossi? Le risposte a queste domande sono state ottenute grazie a un ventaglio di esperimenti sul comportamento e sulla neurofisiologia.

Per determinare il grado di precisione con cui un serpente a sonagli riesce a localizzare un oggetto caldo, Laurence R. Stanford e Michael S. Loop, lavorando con uno di noi (Hartline) all'Università dell'Illinois a Urbana-Champaigne, hanno misurato la precisione dimostrata dai serpenti a sonagli bendati nel colpire. Un serpente a sonagli con gli occhi (ma non le fossette) bendati veniva posto su un piedestallo rialzato al centro di un recinto circolare. Una sorgente di radiazioni infrarosse (la punta riscaldata di un saldatore) veniva poi presentata al serpente ap-

Le strutture del cervello e le vie nervose connesse con il sistema sensoriale infrarosso dei viperidi sono state messe in luce solo di recente. Il problema principale, ora risolto, riguardava come le informazioni raccolte dall'organo a fossetta e inviate come impulsi lungo il trigemino giungessero al tetto ottico, una struttura del mesencefalo che riveste un ruolo fondamentale nell'elaborazione delle informazioni visive.

L'organo a fossetta del serpente a sonagli è formato da una cavità nei tessuti molli e nelle ossa del muso *(a sinistra)*. Una membrana termosensibile, con una superficie di 30 millimetri quadrati e spessa solo 15 micrometri, si estende attraverso la cavità; tale membrana intercetta le radiazioni infrarosse che passano attraverso l'apertura della fossetta.

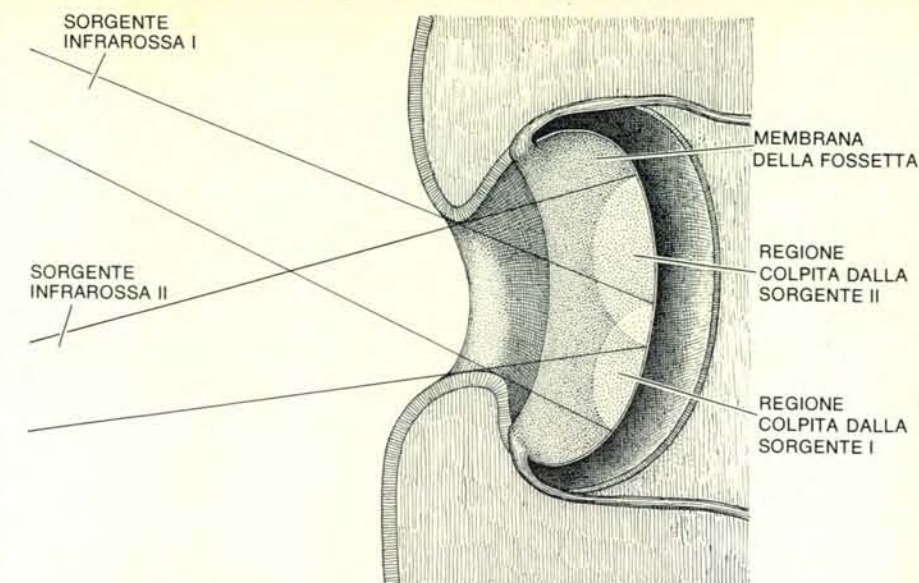
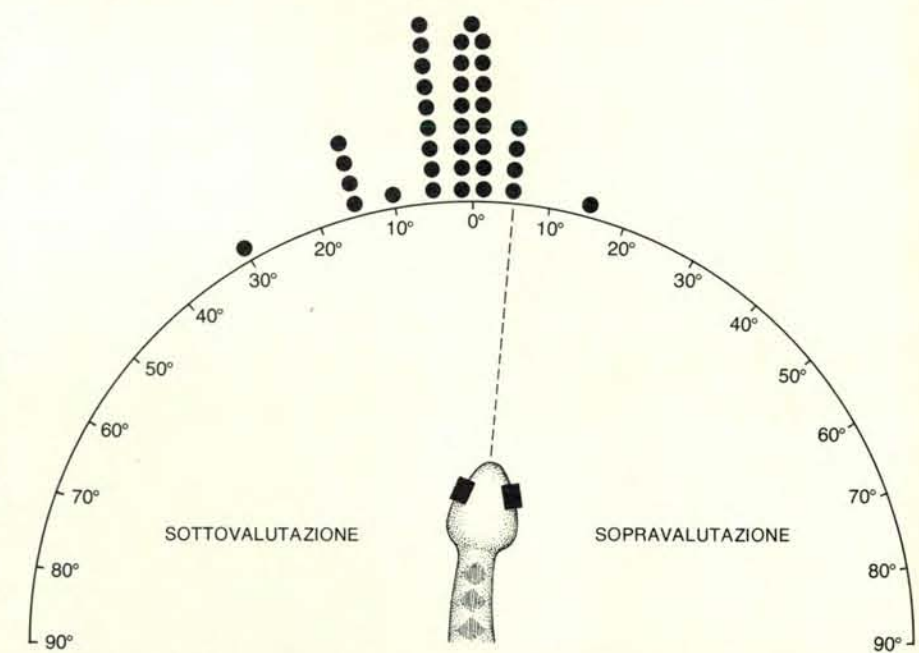
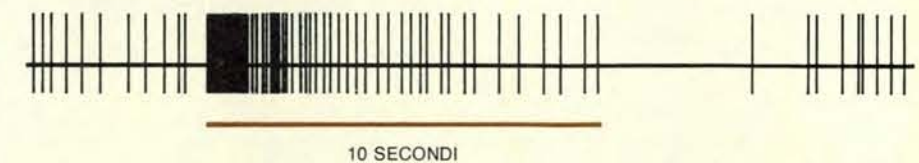
All'interno della membrana *(a destra)* le terminazioni nervose ramificate delle fibre del trigemino si trovano nei primi micrometri della superficie esterna della membrana. Una variazione di soli 0,003 gradi centigradi della temperatura della membrana è sufficiente per elevare il tasso di scarica delle fibre del trigemino.

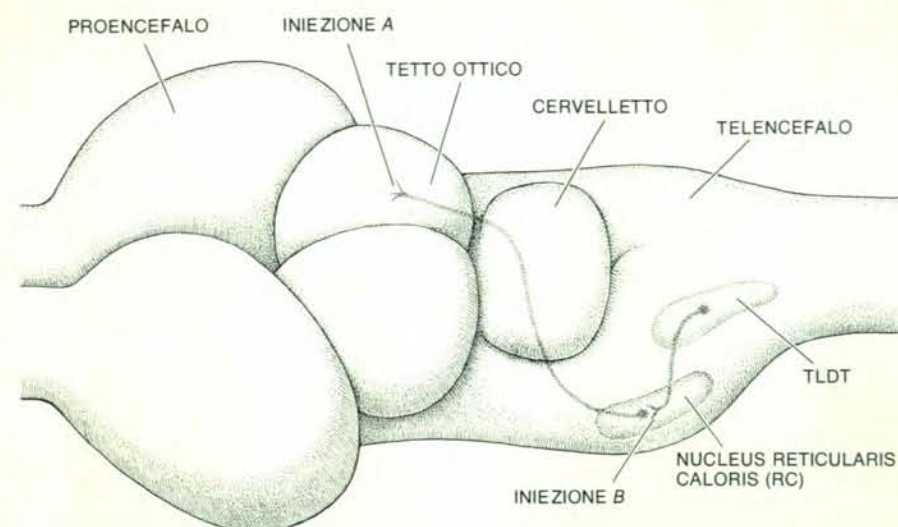
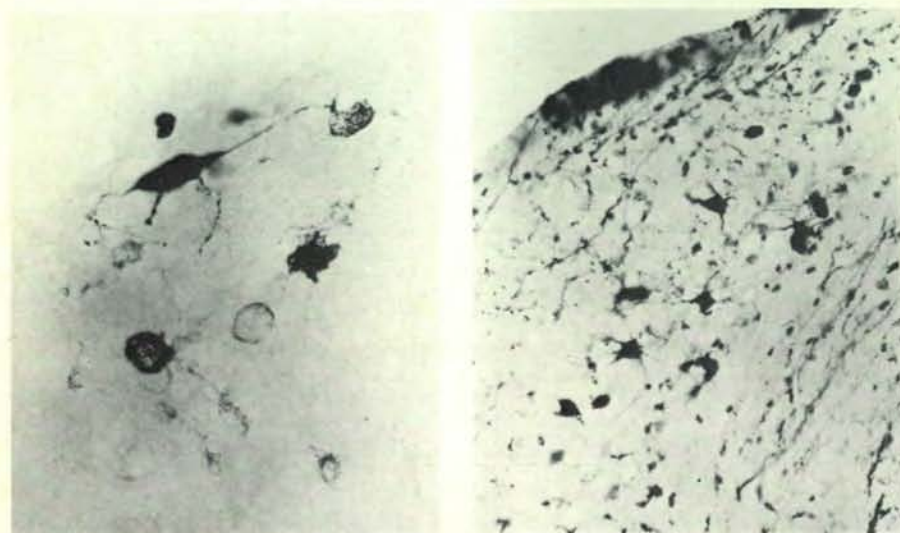
pena al di fuori della sua portata; uno sperimentatore introduceva il saldatore ad angoli variabili da zero a 60 gradi a destra o a sinistra del punto di fronte al serpente. Una telecamera era pronta a riprendere lo scatto atteso; la precisione del colpo sarebbe stata poi misurata esaminando una videoregistrazione.

I serpenti si rifiutavano però di collaborare; essi non colpivano quasi mai il saldatore; apparentemente lo stimolo non era per loro abbastanza significativo. La soluzione, un po' diabolica, è stata quella di far subire al serpente una leggera scossa elettrica poco prima di ogni presentazione del saldatore. I serpenti colpivano, in questo modo, con una certa continuità tale da permetterci di misurare la precisione delle loro risposte. Anche con entrambi gli occhi del tutto coperti, i colpi avevano una tale precisione da cadere entro cinque gradi dal bersaglio immobile, sia che questo si trovasse direttamente di fronte, sia che fosse spostato lateralmente. Questa precisione è piuttosto impressionante, e per un topo sicuramente mortale.

Come può uno studioso misurare la sensibilità all'infrarosso? Il metodo di Bullock consisteva nel registrare la frequenza dei potenziali d'azione, o impulsi nervosi, negli assoni del trigemino che partivano dall'organo a fossetta. Questo metodo neurofisiologico può fornire indicazioni sulla massima sensibilità del recettore della fossetta. Bullock ha trovato che, in assenza di stimolo, singoli assoni portano serie di potenziali d'azione irregolari e a bassa frequenza. Quando la fossetta viene improvvisamente colpita da radiazioni infrarosse moderatamente forti, la frequenza di scarica dei potenziali d'azione aumenta enormemente. Il tasso di scarica cresce entro 100 millisecondi dall'inizio dello stimolo, rimane elevato per breve tempo e diminuisce poi lentamente quasi fino al livello di fondo. Quando lo stimolo cessa, il tasso di scarica scende temporaneamente al di sotto del livello di fondo, poi torna lentamente al livello precedente allo stimolo.

Bullock e i suoi colleghi hanno scoperto che i singoli assoni sono notevolmente sensibili agli stimoli infrarossi. Un oggetto caldo di piccole dimensioni, come la mano di un uomo o un ratto, posti a mezzo metro dalla fossetta, provocano un notevole aumento del tasso di scarica di alcune fibre del trigemino. Un tale stimolo corrisponde a un aumento di circa 0,1 milliwatt per centimetro quadrato dell'energia che incide sull'organo a fossetta. Per dare un termine di paragone, in una giornata invernale serena negli Stati Uniti settentrionali i raggi solari liberano circa 50 milliwatt di radiazione infrarossa per centimetro quadrato.

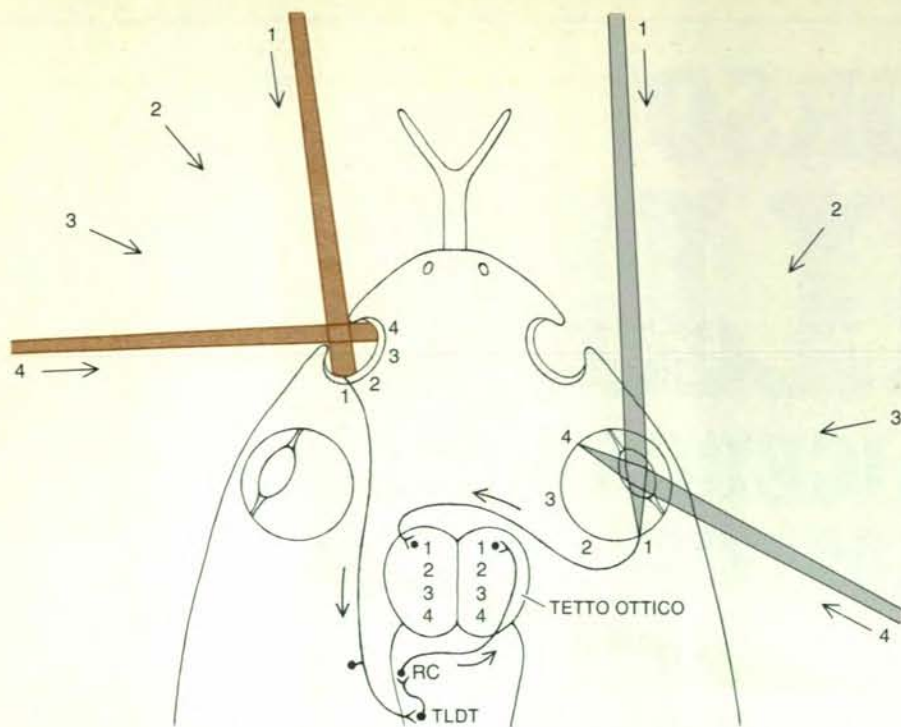
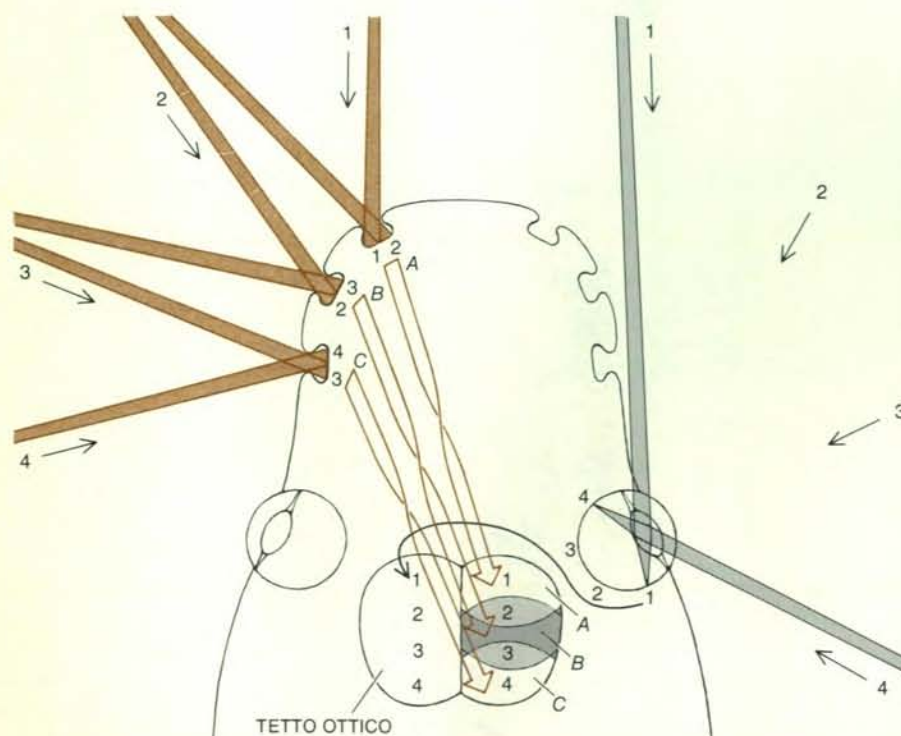
Bullock e F. P. J. Diecke hanno anche effettuato delle misurazioni dirette della sensibilità al calore delle terminazioni delle fibre del trigemino nell'organo a fossetta del serpente a sonagli, dirigendo un flusso di acqua calda nella cavità della fossetta. In questo modo hanno potuto notare che una variazione nella tem-

L'ottica dell'organo a fossetta ricorda uno stenoscopio. Il diametro dell'apertura della fossetta di un serpente a sonagli è di circa due millimetri, cioè più o meno un terzo del diametro della membrana termosensibile. Di conseguenza, la radiazione che proviene da una sorgente infrarossa colpirà solo quella parte di superficie della membrana che cade all'interno di un cono, definito dalla posizione della sorgente e dall'apertura della fossetta. La precisione del sistema viene limitata solamente dalla capacità del serpente nel localizzare i confini della regione colpita.

La precisione spaziale del sistema infrarosso del serpente a sonagli può essere misurata presentando un oggetto caldo a vari angoli alla destra o alla sinistra di un esemplare con gli occhi bendati. I pallini rappresentano un singolo colpo del serpente e mostrano l'errore angolare tra il colpo e il bersaglio. Gli stimoli sono stati presentati fino a 60 gradi a destra e a sinistra dell'animale.

Le fibre nervose del trigemino segnalano le variazioni di temperatura della membrana della fossetta del serpente a sonagli. La linea colorata indica un periodo di prova di 10 secondi durante i quali l'organo a fossetta è stato esposto a radiazioni infrarosse. Le bande verticali sulla linea superiore mostrano gli impulsi registrati in una singola fibra nervosa del trigemino; il tasso di scarica è massimo quando la radiazione ha inizio, ma scende successivamente quasi fino al livello di fondo. Quando la radiazione cessa, si osserva un breve periodo di mancanza di impulsi. Le proprietà delle fibre nervose sensibili all'infrarosso del trigemino sono state scoperte inizialmente da Theodore H. Bullock e dai suoi colleghi dell'Università della California a Los Angeles.

La via anatomica del sistema infrarosso è stata tracciata con l'aiuto di un enzima estratto dal rafano, la perossidasi di rafano (PR), che viene assorbito dalle terminazioni dei neuroni e poi trasportato nei corpi cellulari. Le molecole di PR sono rappresentate dai puntini colorati.

La perossidasi di rafano iniettata nel cervello di un serpente a sonagli è stata seguita colorando delle sezioni sottilissime di tessuto in modo da rivelare la presenza dell'enzima. La PR iniettata nel tetto ottico *(A)* è stata trasportata in un nucleo, o gruppo di corpi cellulari, finora sconosciuto e da noi chiamato *nucleus reticularis caloris* (RC). Iniettando la PR nell'RC *(B)*, sono state trovate alcune cellule colorate nel nucleo TLDT, di cui erano già noti i collegamenti con l'organo a fossetta. TLDT è l'abbreviazione della nuova struttura, detta nucleo del «tratto laterale discendente del trigemino». Gli esperimenti con la PR hanno risolto il problema di come le radiazioni infrarosse che provengono dall'organo a fossetta del serpente a sonagli raggiungano il tetto ottico.

Microfotografie di sezioni del cervello che mostrano come i corpi cellulari marcati con PR vengano evidenziati come profili scuri e irregolari dopo la colorazione. La microfotografia a sinistra è una sezione del nucleo RC fatta dopo l'iniezione di PR nel tetto ottico. La microfotografia a destra rappresenta una sezione del nucleo TLDT dopo che l'enzima è stato iniettato nel nucleo RC. Gli esperimenti sono stati svolti dagli stessi autori e da Edward R. Gruber del Massachusetts Institute of Technology. Al primo esperimento ha partecipato anche E. Kicliter dell'Università di Puerto Rico.

peratura dell'acqua di solo 0,003 gradi centigradi produceva un cambiamento percepibile nel tasso di scarica delle fibre più sensibili.

La notevole sensibilità dell'organo a fossetta nei confronti della radiazione termica dipende dalla sua peculiare struttura: ogni organo a fossetta di un serpente a sonagli o di un boa è formato da una cavità, situata nei tessuti molli e nelle ossa del capo, con un diametro che può variare da uno a cinque millimetri. Nei serpenti a sonagli è presente una sottile membrana di circa 30 millimetri quadrati di superficie che è sospesa nello spazio pieno d'aria della cavità. Questa membrana ha uno spessore di soli 15 micrometri, cioè circa un quarto dello spessore di questa pagina, ed è la componente chiave dell'organo di senso. Sulla superficie della membrana si trovano circa 7000 terminazioni termosensibili degli assoni sensoriali del trigemino. Ognuna di queste terminazioni può eccitare la propria fibra nervosa se viene riscaldata. Terminazioni nervose sensibili al calore più o meno uguali coprono il fondo di ogni fossetta nei boa, e anche se l'epitelio sensibile non si trova sospeso come nei serpenti a sonagli, le terminazioni degli assoni coprono fino a 30 micrometri di superficie.

L'estrema sensibilità al calore degli organi a fossetta è un risultato diretto della vicinanza dei termorecettori alla superficie esterna dell'epitelio sensoriale. Le radiazioni infrarosse che raggiungono la superficie innalzano immediatamente la temperatura dei recettori. La stessa quantità di radiazione che arrivasse sulla pelle nuda di un mammifero innalzerebbe della stessa entità la temperatura dei primi micrometri di pelle, ma i termorecettori dei mammiferi si trovano molto più in profondità. Per stimolarli è necessario riscaldare l'intero spessore del tessuto compreso tra la superficie e i recettori, cioè circa 300 micrometri, e per far ciò in un tessuto spesso 300 micrometri ci vuole un'energia 20 volte superiore a quella necessaria per riscaldare uno strato dello spessore di soli 15 micrometri. La sensibilità dei termorecettori dei mammiferi viene inoltre ridotta dalla perdita di calore per conduzione verso strati ancora più profondi. I serpenti a sonagli evitano questo problema grazie a un'altra caratteristica: dato che la membrana che percepisce il calore è circondata da entrambi i lati dall'aria, essa non perde calore verso tessuti più profondi; l'energia assorbita dalle radiazioni infrarosse riscalda quasi esclusivamente i recettori.

Sebbene nei serpenti a sonagli la sensibilità alla temperatura degli assoni del trigemino sia impressionante, non è straordinaria se la si confronta con le cellule nervose termosensibili di altri animali. La pelle della fronte dell'uomo, ad esempio, ha all'incirca la stessa soglia di sensibilità alla temperatura. La grossa differenza di sensibilità alla radiazione incidente tra i serpenti a sonagli e gli altri animali è soprattutto dovuta alla specia-

Il campo infrarosso e quello visivo sono rappresentati con orientamenti simili sulla superficie del tetto ottico di un serpente a sonagli. Per chiarezza le vie dei due sistemi sensoriali sono mostrate su lati opposti della testa. La superficie anteriore del tetto *(1)* riceve le afferenze dalle zone posteriori della membrana della fossetta e della retina, che controllano approssimativamente le stesse regioni dello spazio. La zona posteriore del tetto ottico *(4)* riceve le afferenze dalle zone anteriori della fossetta e della retina, che «guardano» entrambe verso i lati e dietro all'animale. Benché le proiezioni dei campi di vista infrarosso e visivo sulla superficie del tetto ottico presentino alcune differenze sistematiche, sono abbastanza simili da permettere che ogni regione del tetto ottico riceva informazioni sulla stessa regione dello spazio da entrambe le modalità sensoriali.

Le connessioni nel sistema infrarosso del pitone sono più complesse che nel serpente a sonagli. Gli organi a fossetta del pitone si aprono su zone dello spazio infrarosso differenti, ma che si sovrappongono. Il campo di vista di ogni fossetta viene qui evidenziato sulla regione di proiezione della fossetta sulla superficie del tetto ottico. Le fossette anteriori *(A)* sono rappresentate sulla parte anteriore del tetto, le fossette più arretrate *(C)* sulla parte posteriore. La zona anteriore dell'epitelio sensibile all'infrarosso di ogni fossetta è collegata con la parte posteriore della sua regione di proiezione. Il complesso schema di connessioni dà luogo a una rappresentazione più o meno continua dello spazio infrarosso che corrisponde bene alla rappresentazione visiva.

lizzazione anatomica dell'organo di senso e non è imputabile a differenze fisiologiche nei recettori sensoriali.

Anche le capacità direzionali della sensibilità all'infrarosso possono essere attribuite alle specializzazioni anatomiche della fossetta. Gli organi a fossetta dei serpenti a sonagli hanno aperture grandi meno della metà del diametro della superficie sensoriale; quindi un piccolo oggetto caldo posto a circa mezzo metro di fronte alla fossetta «illuminerà» non più di un quarto dell'intera membrana termosensibile. La posizione della regione illuminata sulla membrana dovrebbe dunque fornire indizi sufficienti per localizzare un oggetto caldo. In breve, l'organo a fossetta funziona in modo analogo a uno stenoscopio.

La precisione con cui il serpente è in grado di localizzare un oggetto caldo deve dipendere dal grado di accuratezza con cui può determinare la posizione della regione colpita sulla membrana. Semplici calcoli trigonometrici dimostrano che per localizzare una piccola sorgente, entro un angolo di cinque gradi, il serpente deve essere in grado di misurare la posizione della macchia calda sulla sua membrana recettiva entro circa 175 micrometri. Con un mosaico di recettori, ognuno dei quali è largo circa 60 micrometri, una tale precisione è abbastanza ragionevole. La diffrazione delle radiazioni infrarosse (le cui lunghezze d'onda dominanti sono comprese tra 6 e 12 micrometri) può offuscare leggermente i confini della macchia calda, ma non in modo tale da poter ridurre in modo significativo la precisione percettiva della fossetta.

Le specializzazioni sviluppate dai serpenti sensibili all'infrarosso non si limitano all'organo a fossetta; viperidi e boidi hanno sviluppato anche una struttura cerebrale unica in modo da elaborare le nuove informazioni raccolte dalle fossette. I primi esperimenti di neurofisiologia condotti sul cervello del serpente a sonagli hanno dimostrato che l'attività di molti neuroni nel tetto ottico del mesencefalo (che in quel periodo si pensava fosse deputato in modo privilegiato alla visione) era controllata da stimoli infrarossi. Evidentemente il sistema infrarosso, parallelamente alla sua evoluzione da senso collegato alla pelle del capo a senso simile alla vista, ha sviluppato una cospicua proiezione nel tetto. Questa scoperta, fatta indipendentemente da uno di noi (Hartline), che in quel momento lavorava all'Università della California a San Diego, e da S.-I. Terashima e R. C. Goris della Università medico-dentistica di Tokyo, presentava un interessante problema anatomico. In altri rettili e nei mammiferi il trigemino termina nella regione trigeminale del telencefalo; nel serpente a sonagli come fanno le informazioni che hanno origine nel trigemino a giungere al tetto ottico del mesencefalo?

Su questo problema hanno iniziato a lavorare due gruppi di ricercatori. Dolores M. Schroeder e Loop, che lavo-

rano alla University of Virginia Medical School, hanno usato degli ioni cobalto per individuare i collegamenti del trigemino del serpente a sonagli, ponendo l'estremità recisa di una ramificazione del nervo che riveste la fossetta in una soluzione di cloruro di cobalto e facendo passare una corrente elettrica tra la soluzione e il cervello. La corrente spingeva gli ioni cobalto lungo gli assoni del nervo fino alle terminazioni degli assoni stessi. Quando Schroeder e Loop hanno trattato con un colorante specifico per il cobalto sezioni di cervello, hanno scoperto che le fibre del trigemino, invece di terminare nella normale area trigeminale del telencefalo, finivano in un nuovo nucleo, oggi chiamato TLDT (abbreviazione di nucleo del «tratto laterale discendente del trigemino»).

G. J. Molenaar dell'Università di Leida ha identificato indipendentemente nel pitone, un serpente della famiglia dei boidi, un nucleo che quasi certamente è lo stesso. Sembra che il nucleo TLDT sia caratteristico dei serpenti che presentano una sensibilità specializzata per l'infrarosso e non è stato trovato in alcun altro animale; pare inoltre che abbia a che fare esclusivamente con il sistema sensoriale infrarosso.

La nostra curiosità era stata stimolata: se gli impulsi nervosi portati dagli assoni del trigemino giungono al nucleo TLDT, tramite quale via le informazioni sensoriali infrarosse arrivano al tetto ottico? Un modo per individuare i collegamenti all'interno del cervello consiste nell'usare l'enzima perossidasi estratto dal rafano. La molecola di perossidasi di rafano, di grandi dimensioni, viene assorbita in modo specifico dalle terminazioni degli assoni di singoli neuroni e viene trasportata in altre regioni delle stesse cellule. Una tecnica di colorazione rende la perossidasi di rafano chiaramente visibile in sezioni di cervello e rivela la posizione dei corpi cellulari dei neuroni che hanno assorbito la perossidasi.

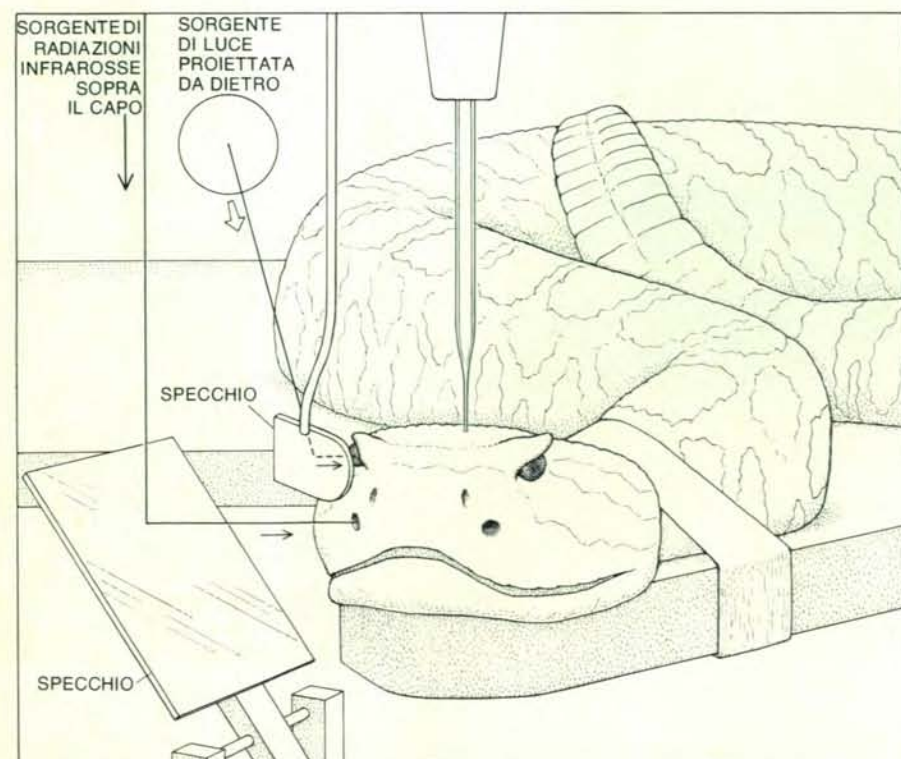
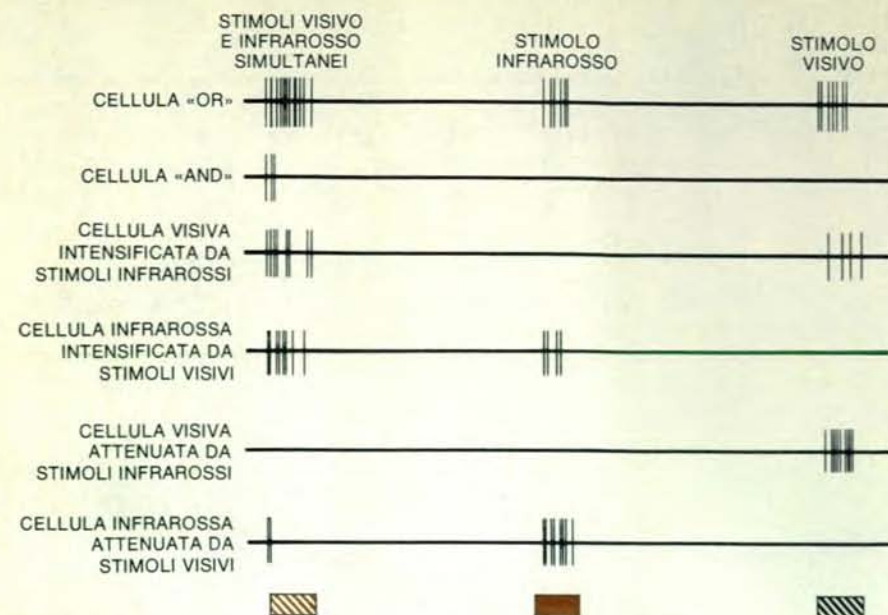
Lavorando in collaborazione con Edward R. Gruberg del Massachusetts Institute of Technology ed E. Kicliter dell'Università di Puerto Rico, abbiamo iniettato della perossidasi di rafano in profondità nel tetto ottico del serpente a sonagli, dove si trovano le afferenze sensoriali infrarosse. Se le cellule del TLDT si proiettassero direttamente nel tetto ottico, pensavamo, le loro terminazioni assoniche nel tetto preleverebbero la perossidasi e la ritrasporterebbero nel TLDT. Con una certa sorpresa, invece, non abbiamo trovato neuroni colorati nel TLDT, mentre abbiamo trovato un gruppo di cellule notevolmente colorate in una regione differente del telencefalo. Abbiamo allora ipotizzato che questo gruppo di cellule, che abbiamo chiamato *nucleus reticularis caloris* (RC), funzioni da ripetitore tra il TLDT e il tetto.

A questo punto dovevamo determinare se il nostro nucleo era o meno collegato direttamente con il TLDT. Abbiamo studiato questo problema con Gruberg iniettando perossidasi di rafano nel nucleo RC. Questa volta abbiamo trovato cellule colorate nel TLDT. Stanford e Schroeder hanno studiato a fondo questo stesso problema utilizzando una tecnica basata sulla degenerazione delle cellule nervose. Distruggendo selettivamente le cellule del TLDT del serpente a sonagli e sfruttando una procedura di colorazione che rivelava gli assoni degenerati di queste cellule, essi sono riusciti a individuare gli assoni che portano dal TLDT al RC.

Dunque entrambe le metodologie sperimentali hanno fornito la stessa risposta: le cellule del nucleo di TLDT si proiettano direttamente nel nucleo RC. Anche se nessuno di questi esperimenti avrebbe potuto chiarire i ruoli funzionali dei nuclei TLDT e RC nel sistema sensoriale infrarosso, essi hanno almeno spiegato il modo in cui le informazioni infrarosse raggiungono il tetto ottico. Nel serpente a sonagli i nuclei TLDT e RC si sono evoluti per elaborare le informazioni sensoriali infrarosse e trasmetterle al tetto ottico.

L'efficacia del sistema infrarosso nel guidare il colpo del serpente dimostra chiaramente che tale sistema è in grado di localizzare gli oggetti nello spazio con notevole precisione. In questo ruolo esso funziona in modo estremamente simile al sistema visivo di altri animali. Gli studi in campo neurofisiologico nostri e di colleghi hanno rivelato numerose somiglianze tra i meccanismi che permettono ai sistemi visivo e infrarosso di localizzare gli oggetti.

Bullock, e in seguito J. A. DeSalvo, insieme a uno di noi (Hartline), hanno descritto in che modo i singoli assoni del trigemino rispondono a stimoli termici, che si trovano in una qualsiasi posizione all'interno di un campo visivo descritto da un cono il cui vertice si trova nella fossetta. Questi coni visivi misurano in genere tra 45 e 60 gradi di ampiezza e vengono chiamati, per analogia con i loro corrispondenti visivi nella retina, campi recettivi eccitatori. La maggior parte delle fibre del trigemino si comportano, dunque, come se i loro campi recettivi fossero limitati semplicemente dalle ombre prodotte dal bordo della fossetta. La risposta di ciascun neurone del trigemino non comunicherebbe con precisione la posizione di uno stimolo, ma una popolazione di molti neuroni, la cui attività potrebbe essere analizzata dal cervello, sarebbe in grado di farlo. In una analisi di questo tipo il tetto ottico svolge probabilmente un ruolo importante.

Il tetto ottico, conosciuto anche nei mammiferi come collicolo superiore, è uno dei centri visivi più importanti del cervello. È forse il centro principale per l'elaborazione delle informazioni sulla localizzazione spaziale degli oggetti. Nelle scimmie, il tetto permette di dare l'av-

Il senso infrarosso e quello visivo del serpente a sonagli sono stati stimolati con speciali attrezzature nel laboratorio degli autori dell'Eye Research Institute della Retina Foundation a Boston. Il microelettrodo collocato sul capo del serpente registrava l'attività di singole cellule del tetto ottico. Uno specchio girevole rifletteva nell'organo a fossetta di destra le radiazioni infrarosse provenienti da una sorgente posta direttamente al di sopra del serpente, mentre le immagini visive proiettate da dietro su uno schermo venivano dirette nell'occhio destro da uno specchio di dimensioni minori. Lo stimolo visivo era costituito da una macchia circolare luminosa.

Sono state identificate sei classi di cellule del tetto ottico che rispondono a stimoli infrarossi e visivi. I rettangoli in basso mostrano il tempo di presentazione di stimoli combinati infrarossi e visivi *(tratteggi bianchi e rossi)*, di stimoli infrarossi da soli *(rosso)* e di stimoli visivi da soli *(tratteggi bianchi e neri)*. Le cellule del tetto denominate «or», «and» e «intensificate» mostrano interazioni eccitatorie delle afferenze infrarosse e visive mentre le due classi «attenuate» di cellule mostrano delle interazioni inibitorie se le due afferenze vengono presentate insieme.

vio e di controllare la direzione dello sguardo. Nelle rane e nei rospi probabilmente dà origine e presiede alla cattura di prede volanti e non.

Nel tetto ottico le informazioni visive sono organizzate in modo ordinato: i collegamenti tra la retina e il tetto (uniti dal nervo ottico) sono fatti in modo tale da formare sulla superficie del tetto una mappa dello spazio visivo. Proprio come una carta stradale rappresenta l'organizzazione spaziale delle strade sulla superficie terrestre, così la mappa visiva sulla superficie del tetto rappresenta lo spazio come viene percepito dall'occhio. Ad esempio, la parte anteriore del tetto rappresenta la porzione di spazio visivo che sta di fronte all'animale; la parte posteriore del tetto rappresenta lo spazio verso i lati e sul retro. La rappresentazione ordinata dello spazio visivo sulla superficie del tetto ottico è una caratteristica costante dei vertebrati, dai pesci ai primati.

La vista non è l'unico senso in grado di comunicare al cervello dove si trovano gli oggetti, poiché anche l'udito fornisce importanti indizi sulla posizione, e così il tatto. Anche questi due sensi sono rappresentati nel tetto ottico dei mammiferi; inoltre sia il sistema acustico che il sistema sensoriale somatico vengono rappresentati nel tetto in un modo spaziale ordinato che coincide grosso modo con la mappa del tetto del sistema visivo. La parte anteriore del tetto rappresenta la regione frontale dello spazio acustico: i suoni provenienti dal davanti eccitano di preferenza questa regione del tetto. Analogamente, la parte anteriore del tetto risponde generalmente a stimolazioni tattili della parte anteriore del corpo; ad esempio del muso, di un animale.

Il sistema infrarosso è organizzato spazialmente in modo analogo nel tetto ottico del serpente a sonagli? Leonard J. Kass, Loop e uno di noi (Hartline) hanno studiato questo problema all'Università dell'Illinois servendosi di tecniche neurofisiologiche. Abbiamo registrato le risposte elettriche delle cellule del tetto da una griglia di posizioni sulla sua superficie, e abbiamo determinato in quali posizioni dello spazio uno stimolo infrarosso provocava una risposta massima dei neuroni.

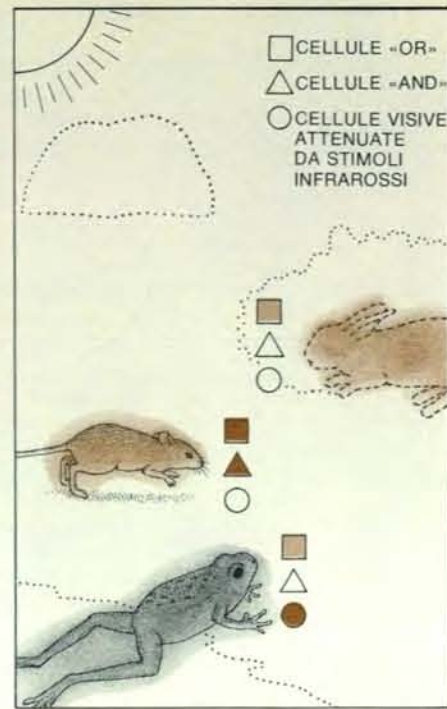
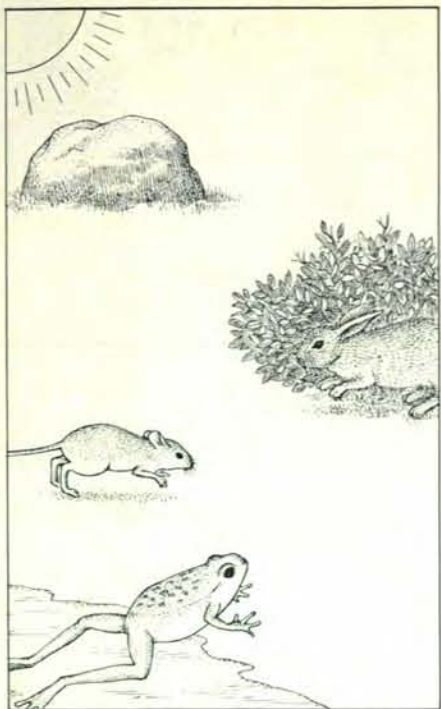
Dopo avere analizzato i dati che si riferivano a vari serpenti, siamo stati ricompensati dalla scoperta che non solo la sensibilità all'infrarosso è organizzata spazialmente nel tetto ottico, ma anche che la mappa infrarossa del tetto segue le stesse regole di quella visiva. La parte anteriore del tetto risponde agli stimoli infrarossi che provengono dal davanti; la parte posteriore agli stimoli provenienti dai lati. In breve, la mappa visiva e quella infrarossa sono, almeno in prima approssimazione, in registro. Come vedremo, la similarità delle organizzazioni spaziali dei sistemi visivo e infrarosso offre dei vantaggi nell'elaborazione dei segnali provenienti dalle due modalità sensoriali.

L'orientamento della mappa dello spazio infrarosso sul tetto ottico rappresenta la caratteristica che più si allontana dallo schema generale di connessioni del sistema sensoriale somatico dei vertebrati, anche se il sistema infrarosso si è presumibilmente evoluto da un senso somatico. Nella maggior parte degli altri vertebrati la parte anteriore del tetto rappresenta la parte anteriore della superficie corporea. Nel serpente a sonagli, tuttavia, la parte anteriore del tetto rappresenta la parte della membrana della fossetta più lontana dal muso dell'animale invece che la parte più vicina. Questa inversione delle connessioni mantiene nel tetto ottico la normale rappresentazione spaziale, ma non quella della superficie corporea.

I boidi presentano un esempio di connessioni tra pelle e tetto ancora più complesso. Il pitone reticolato, ad esempio, ha 13 fossette su ciascun lato del capo; ogni fossetta è sostanzialmente un sistema ottico a inversione, che necessita di una mappa rovesciata sul tetto, come quella del serpente a sonagli. Ci si potrebbe aspettare che le fossette anteriori fossero collegate con la parte anteriore del tetto, secondo il solito schema di non inversione dei vertebrati. Come può questa schiera di fossette essere collegata con il tetto in modo da formare una verosimile rappresentazione dello spazio infrarosso? E. C. Haseltine dell'Università dell'Indiana ha affrontato questo affascinante problema in collaborazione con uno di noi (Hartline).

Lo schema di connessioni somatiche che abbiamo trovato è il più complesso che sia stato osservato fino a oggi in qualsiasi animale. Per cominciare abbiamo localizzato la regione del tetto ottico che riceve le connessioni dall'intera superficie sensoriale di una singola fossetta, cioè la regione di proiezione della fossetta. Quando abbiamo riportato in mappa le connessioni della fossetta all'interno di questa regione, abbiamo scoperto che la parte posteriore della superficie sensoriale di ogni fossetta corrispondeva alla parte anteriore della sua area di proiezione nel tetto. La mappa di ciascuna fossetta risulta, quindi, rovesciata rispetto alle mappe somatiche normali dei vertebrati. Mettendo a confronto le regioni di proiezione di varie fossette abbiamo però trovato che il solito schema di mappatura dei vertebrati corrisponde al vero: le fossette anteriori proiettano verso la parte anteriore del tetto e le fossette posteriori proiettano verso la parte posteriore del tetto. Vi è poi un ultimo aspetto complesso: fossette adiacenti, i cui campi di percezione infrarossa si sovrappongono ampiamente, hanno anche regioni di proiezione nel tetto che si sovrappongono abbondantemente.

Il risultato finale è una singola rappresentazione coerente dello spazio infrarosso sulla superficie del tetto ottico del pitone. È molto importante che la mappa visiva e quella infrarossa nel tetto del pitone corrispondano adeguatamente l'una con l'altra; evidentemente è andata evolvendosi una intricata ridisposizione delle connessioni tra il sistema sensoriale somatico e il tetto in modo da mantenere certi rapporti spaziali tra le due modalità sensoriali.

La corrispondenza tra rappresentazioni visiva e infrarossa del mondo nel tetto ottico del serpente a sonagli e del pitone suggerisce un'affascinante possibilità: probabilmente il tetto mette a confronto o combina i segnali provenienti dai due sensi. Già sapevamo che nei mammiferi

In queste illustrazioni sono rappresentati, in maniera schematica, i modi ipotetici in cui un serpente a sonagli percepisce il mondo. La figura a sinistra rappresenta una scena come potrebbe essere registrata dal solo sistema visivo. La figura al centro è la stessa scena come potrebbe essere percepita dal solo sistema infrarosso del serpente. Unicamente gli oggetti caldi *(in colore)* o freddi *(in grigio)* e in movimento possono essere rivelati; gli oggetti immobili come il cespuglio e lo stagno non possono venire rappresentati perché non stimolano a sufficienza il sistema infrarosso. Anche la roccia riscaldata dal Sole è assente dall'immagine infrarossa perché non è in movimento. Le immagini potrebbero essere sfocate a causa delle scarse proprietà rappresentative dell'organo a fossetta. La figura a destra suggerisce in che modo il tetto ottico analizzerebbe le afferenze combinate dei due sistemi sensoriali. I triangoli, i quadrati e i cerchi rappresentano le risposte di tre delle sei classi di cellule bimodali del tetto. Le cellule «or» verrebbero intensamente attivate dal topo caldo e chiaramente visibile *(in colore scuro)* e attivate in modo meno intenso *(in colore chiaro)* dalla componente infrarossa del coniglio, difficile da vedere contro il cespuglio, e dalla componente visiva della rana, che è fredda. Le cellule «and» verrebbero attivate solo dal topo, che è sia caldo sia chiaramente visibile, mentre la pelle fredda della rana servirebbe a disinibire le cellule visive «attenuate» da stimoli infrarossi provocando una loro intensa risposta.

alcuni neuroni del tetto rispondono a stimoli sia visivi sia acustici o a stimoli sia visivi sia tattili. Abbiamo dunque ipotizzato che anche nel tetto del serpente a sonagli possano esistere neuroni «multimodali» di questo tipo.

Non siamo rimasti delusi: infatti i primi neuroni multimodali nel tetto di un serpente sono stati identificati da uno di noi (Hartline) e da Kass e Loop. Gli stimoli infrarossi venivano forniti da una mano o da un saldatore, gli stimoli visivi da una striscia di cartone termicamente neutra. Abbiamo scoperto che alcuni neuroni del tetto ottico rispondono agli oggetti caldi sia alla luce che al buio totale. Gli stessi neuroni rispondono alla striscia di cartone, ma solo se le luci della stanza sono accese. Abbiamo chiamato queste cellule neuroni «or», dal momento che possono venire eccitate per mezzo della stimolazione di una o dell'altra delle due modalità.

Abbiamo anche trovato alcuni neuroni del tetto che rispondono a oggetti caldi, ma solo se la stanza è illuminata. Questi neuroni non rispondono a una striscia di cartone termicamente neutra neppure se la stanza è abbondantemente illuminata, né rispondono a oggetti caldi in una stanza buia. Li abbiamo chiamati neuroni «and», dato che sembra che per la loro risposta siano necessari stimoli sia visivi sia infrarossi. Neuroni come questi non sono mai stati descritti per nessuna combinazione di modalità sensoriali in altri animali.

Noi due (Newman e Hartline) ci siamo resi conto che avremmo potuto trovare altri tipi di interazioni tra la modalità visiva e infrarossa se avessimo ideato esperimenti specifici in grado di porle in evidenza. Abbiamo costruito uno stimolatore controllato dal calcolatore che potesse presentare stimoli infrarossi e visivi le cui intensità, posizione e durata fossero fra loro indipendenti. Come in alcuni trucchi degli illusionisti, è stato possibile ottenere questo effetto per mezzo di alcuni specchi ben piazzati *(si veda l'illustrazione a pagina 120)*.

In un tipico esperimento abbiamo inserito un microelettrodo da registrazione nel tetto ottico di un serpente a sonagli in modo tale da poter registrare le risposte elettriche di un singolo neurone. Il nostro calcolatore ha quindi generato una sequenza di stimoli infrarossi, visivi e infrarossi-visivi combinati e ha registrato le risposte della cellula. Abbiamo successivamente spostato il microelettrodo fino a incontrare un'altra cellula e in questo modo abbiamo campionato e saggiato i risultati di numerosi neuroni.

I nostri sospetti a proposito delle insolite modalità di interazione si sono dimostrati fondati: abbiamo, infatti, trovato varie e nuove interazioni tra la percezione dell'infrarosso e quella visiva. Le risposte delle cellule del tetto, e quindi le cellule stesse, si potevano dividere in sei classi, due delle quali sono rappresentate rispettivamente dai neuroni «or» e «and». Abbiamo avuto diverse conferme che le cellule «or» rispondono tanto a stimoli infrarossi e visivi presentati da soli quanto alle combinazioni di stimoli infrarossi e visivi, mentre le cellule «and» rispondono solo alla presentazione simultanea di stimoli infrarossi e visivi.

I neuroni del tetto di altre due classi mostrano delle interazioni «intensificanti» delle modalità visive e infrarosse. Abbiamo definito queste cellule come «infrarosse intensificate da stimoli visivi» oppure «visive intensificate da stimoli infrarossi», a seconda dello stimolo che compie l'intensificazione. Ad esempio, una cellula «infrarossa intensificata da stimoli visivi» mostrerà una risposta moderata (diciamo quattro impulsi nervosi) a uno stimolo infrarosso presentato da solo e nessuna risposta a uno stimolo visivo; presentando, invece, stimoli visivi e infrarossi insieme si avrà una risposta più marcata (otto impulsi). L'influenza eccitatoria dello stimolo visivo, anche se lo stimolo non è abbastanza forte da generare da solo una risposta, può intensificare in modo significativo la risposta della cellula allo stimolo infrarosso. Analogamente, una cellula «visiva intensificata da stimoli infrarossi» non mostrerà alcuna risposta a uno stimolo infrarosso, ma darà una risposta intensificata a uno stimolo visivo quando questo sarà accompagnato da uno stimolo infrarosso.

I neuroni del tetto delle ultime due classi presi in considerazione mostrano delle interazioni inibitorie delle modalità infrarossa e visiva. I neuroni di una classe, da noi chiamati cellule «visive attenuate da stimoli infrarossi», rispondono vigorosamente a uno stimolo visivo presentato da solo, ma non rispondono o rispondono debolmente a uno stimolo combinato visivo e infrarosso. L'influenza inibitoria dello stimolo infrarosso riduce o sopprime completamente la risposta a quello visivo. D'altra parte, con cellule di questa classe, la rimozione di un oggetto caldo o l'introduzione di uno freddo intensificherà le risposte visive. I neuroni dell'altra classe, in cui una risposta all'infrarosso viene inibita da uno stimolo visivo, sono stati chiamati da noi cellule «infrarosse attenuate da stimoli visivi».

Quale potrebbe essere il ruolo funzionale svolto dai neuroni infrarossi-visivi da noi identificati? Cerchiamo di rispondere a questa domanda considerando come potrebbe funzionare il tetto quando un serpente a sonagli è in cerca di preda. Una descrizione semplificata di questo comportamento potrebbe portare a dividerlo in vari processi: la rilevazione e la localizzazione dello stimolo, l'identificazione dello stimolo e la scelta del movimento di orientamento appropriato; può poi essere coinvolto anche un processo in cui l'attenzione dell'animale si focalizza in modo preferenziale su una regione dello spazio; solo uno stimolo in quella regione provocherebbe un movimento del capo o degli occhi.

Le cellule «or» e le cellule «intensificate» sarebbero utili nel processo di rilevazione: le cellule «or» segnalerebbero gli eventi in una particolare regione dello spazio, sia che gli eventi venissero percepiti per mezzo delle modalità visiva o infrarossa sia che lo fossero per mezzo di entrambe le modalità insieme. Le cellule «or» e quelle «intensificate» rispondono in modo più intenso agli stimoli combinati che agli stimoli evocati dalle radiazioni visibile o infrarossa da sole. La rilevazione e la localizzazione verrebbero migliorate da tali interazioni eccitatorie, in particolare quando nessuna delle due modalità viene intensamente eccitata. Questo avrebbe luogo per esempio se nella debole luce del crepuscolo un topo si trovasse appena oltre la massima distanza a cui può essere percepito dall'organo a fossetta. La somma delle modalità verrebbe effettuata nella corretta posizione del tetto ottico grazie all'organizzazione spaziale simile delle modalità visiva e infrarossa.

Può darsi che i neuroni «or» e «intensificati» abbiano a che fare anche con un processo di attenzione, come nel caso in cui un oggetto caldo e visibile posto lateralmente al serpente attirasse l'«attenzione» del dispositivo di orientamento del cervello. La componente infrarossa dello stimolo genererebbe un elevato stato eccitatorio tra le cellule «or» e le unità intensificate in una estesa regione del tetto, poiché i campi recettivi all'infrarosso sono piuttosto ampi. Una tale elevata eccitazione potrebbe preparare quella regione del tetto per una stimolazione visiva da parte dello stesso oggetto o di uno vicino. Lo stimolo infrarosso avrebbe dunque la funzione di attirare l'attenzione verso la parte corrispondente dello spazio visivo.

Può darsi che le interazioni infrarosso-visivo a livello del tetto contribuiscano anche al processo di identificazione stimolando dei neuroni di «rilevazione di caratteristiche» che rispondono solo a specifiche combinazioni di stimoli infrarossi e visivi. Le cellule «and», per esempio, vengono attivate solo da oggetti che possono venire percepiti simultaneamente dall'organo a fossetta e dall'occhio: tali stimoli simultanei fanno sì che i neuroni «and» emettano una breve scarica di impulsi ad alta frequenza. Abbiamo trovato che la maggior parte degli stimoli più efficaci per queste cellule è costituita da oggetti di piccole dimensioni, caldi e in movimento; con un po' di fantasia questi neuroni potrebbero essere descritti come «rivelatori di topi».

Le cellule visive attenuate dall'infrarosso, d'altra parte, rispondono al meglio a oggetti visivi termicamente neutri o freddi: foglie che volano trasportate dal vento o anche una rana che salta sulla riva di uno stagno. L'attività visiva di queste cellule verrebbe diminuita dalle radiazioni infrarosse degli animali a sangue caldo. In questi esempi le interazioni tra le due modalità sono essenziali per una corretta identificazione dello stimolo.

Le interazioni tra le due modalità da noi descritte dimostrano che nel tetto ottico del serpente a sonagli le modalità sensoriali infrarosse-visive vengono integrate in modo significativo. È possibile che anche nel tetto ottico di altre specie operino simili interazioni che combinano varie modalità? È possibile, per esempio, che nei mammiferi gli stimoli tattili e uditivi vengano combinati in qualche modo con gli impulsi visivi? Non lo sappiamo ancora poiché non sono stati effettuati esperimenti che permettano di identificare tali interazioni. Siamo, però, pronti a scommettere che i tipi di interazioni tra le modalità da noi osservate nei serpenti siano presenti e funzionalmente importanti nel tetto ottico della maggior parte degli altri vertebrati.

È strano che si sappia così poco del modo in cui i serpenti fanno uso del senso dell'infrarosso nel loro habitat naturale; nessuno sa se i serpenti a sonagli e i pitoni facciano uso dei rivelatori a infrarossi per cacciare le prede di giorno, quando funziona anche il sistema visivo, né è noto se i serpenti facciano uso del sistema infrarosso per identificare i predatori e per trovare un luogo di riposo caldo e confortevole.

Cocaina

Quando un consumatore assume questa droga in condizioni controllate spesso non riesce a distinguerla da altre droghe o da un placebo. Ciò nonostante, per gli alti profitti in gioco, essa è oggetto di vasti traffici

di Craig Van Dyke e Robert Byck

Pochi centigrammi di cloridrato di cocaina, triturato finemente e disposto su una superficie liscia lungo tante righe (come filari di polvere), possono essere aspirati con il naso in pochi secondi con l'aiuto di un pezzetto di carta arrotolato. L'inalazione dà luogo rapidamente a sensazioni euforiche e a un'impressione di chiarezza e di energia della mente, sensazioni che generalmente durano mezz'ora. Sebbene il crescente interesse per la cocaina negli Stati Uniti e nell'Europa occidentale sia in parte legato a questa semplice esperienza edonistica, la reale importanza di questa droga è dovuta all'interazione di fattori sociali, economici e politici: l'attuale status della cocaina quale droga alla moda per uso «voluttuario», il suo prezzo artificialmente gonfiato, il controllo della sua distribuzione da parte di organizzazioni criminali, le tensioni che il suo spaccio attraverso baratti e commerci, i quali sfuggono a ogni tassazione, crea nel sistema monetario e in quello fiscale, la minaccia alle libertà civili rappresentata da leggi antidroga non fondate su una razionale valutazione degli effetti della droga stessa, infine il potenziale di corruzione dei pubblici funzionari e di discredito della legge insito nel commercio della cocaina.

Le stime del costo e del consumo di cocaina sono influenzate dai preconcetti delle fonti di informazione e perciò devono essere considerate con scetticismo. Secondo un rapporto presentato nel 1979 dal White House Strategy Council on Drug Abuse, circa 10 milioni di americani avevano assunto cocaina durante i 12 mesi precedenti, rispetto ai 10 000 di vent'anni prima. Il National Narcotics Intelligence Consumers Committee (NNICC) ha calcolato che nel 1979 sono entrati illegalmente negli Stati Uniti tra 25 000 e 31 000 chilogrammi di cocaina. L'importazione di cocaina negli Stati Uniti nel corso del 1980 è stata stimata tra 40 000 e 48 000 chilogrammi. Queste stime quantitative sono state dedotte dalla valutazione della capacità di raffinazione dei laboratori clandestini, dalla stima delle quantità di prodotto raffinato che sono state piazzate sul mercato interno ed estero, infine dalla stima dei redditi da produzione delle foglie di coca greggia nei paesi di origine. Le stime dei redditi da coltivazione, elaborate nei paesi di origine, in molti casi (e forse ciò non è sorprendente), sono inferiori di due o più volte a quelle elaborate negli Stati Uniti.

Lo spacciatore clandestino al dettaglio riesce a vendere per una somma tra i 100 e i 140 dollari a grammo una sostanza bianca, cristallina, costituita fra il 10 e l'85 per cento da cocaina e da varie altre sostanze. Il NNICC, sulla base delle proprie stime relative alle importazioni di cocaina nel 1980, valuta il fatturato dello spaccio al dettaglio tra 27 e 32 miliardi di dollari. Se il commercio della cocaina venisse incluso da «Fortune» nella sua lista delle 500 maggiori imprese industriali, la cocaina dovrebbe essere classificata al settimo posto per volume di vendite sul mercato interno, tra la Ford Motor Company e la Gulf Oil Corporation. Sulla base delle stime americane, il valore monetario dell'esportazione di cocaina boliviana deve essere ritenuto attualmente superiore a quello dell'industria legale più diffusa in quel paese, lo stagno. L'esportazione della cocaina più raffinata della Colombia raggiunge un valore annuale di 1 miliardo di dollari, cioè la metà del valore della coltivazione del caffè.

È fin troppo facile supporre che le conseguenze fisiologiche e sociali dell'uso di cocaina siano correlate con la sua popolarità e con la sua importanza economica. In realtà soltanto recentemente è cominciato il lavoro di accertamento delle implicazioni mediche e psicologiche dell'uso di cocaina a breve e a lungo termine. Nel 1975 il National Institute on Drug Abuse ha dato inizio a un progetto di ricerca volto a stabilire la farmacologia della cocaina nell'uomo. Da allora molti studiosi sono stati impegnati in ricerche su vari aspetti della cocaina, tra cui una dettagliata descrizione dei suoi effetti, la sua distribuzione e il suo metabolismo nell'uomo, le sue proprietà di rinforzo e la sua capacità potenziale di provocare abuso.

Dopo sei anni di lavoro attorno a questo problema il nostro gruppo presso la Yale University School of Medicine è stato in grado di descrivere con ragionevole sicurezza l'andamento dei fondamentali effetti farmacologici conseguenti all'introduzione di cocaina nel corpo umano attraverso varie vie. Siamo stati altresì in grado di associare l'andamento della concentrazione di cocaina nel sangue con misure ripetibili dei suoi effetti psicologici. Nonostante la cocaina risulti intensamente piacevole per certe persone, abbiamo trovato che la sua capacità di provocare uno specifico stato è stata sopravvalutata: i nostri soggetti, tutti sperimentati consumatori di cocaina, non sono stati capaci di distinguere tra una singola dose di cocaina, introdotta per via nasale, e una stessa quantità di lidocaina, anestetico locale sintetico. Ricercatori dell'University of Chicago School of Medicine hanno constatato che i loro soggetti non erano in grado di distinguere gli effetti immediati dell'iniezione endovenosa di cocaina da quelli dell'amfetamina, nonostante siano del tutto evidenti, in fin dei conti, le differenze tra le due droghe. Questi risultati sono i primi passi per arrivare a separare l'esaltante alone mitologico che circonda la cocaina da un'informazione attendibile circa i suoi effetti.

Nello stesso periodo di tempo c'è stata una rinascita dell'interesse sia per la storia del consumo di cocaina sia per il suo significato nella cultura india della regione delle Ande che comprende parti degli attuali stati di Colombia, Ecuador, Perù, Bolivia e Cile. Naturalmente l'esperienza sociale e farmacologica delle foglie di coca, tipica della regione delle Ande, non può essere immediatamente paragonata con la comune esperienza sociale della ben più potente sostanza che è il cloridrato di cocaina, ma può essere utile per acquisire una prospettiva culturale della droga. Tale prospettiva è, per una ben informata condotta di pubblici poteri, almeno altrettanto importante della conoscenza dell'azione biochimica e farmacologica della cocaina.

La masticazione delle foglie di coca, rappresentata in questa ceramica, ha avuto un ruolo importante nella cultura dei popoli delle Ande durante almeno 5000 anni. È appunto dalle foglie di coca che viene estratto l'alcaloide cocaina; i masticatori di foglie assorbono la sostanza direttamente. La scultura mostra un uomo la cui gota sinistra è rigonfia di foglie di coca e che tiene in mano gli arnesi utilizzati per la masticazione. La piccola zucca nella mano sinistra contiene *llicta*, una polvere alcalina che egli aggiunge alla coca che ha in bocca inumidendo con la saliva la spatola che tiene nella mano destra. In questo modo la polvere aderisce alla spatola. Nel 1979 David Paly della School of Medicine della Yale University ha dimostrato che l'aggiunta di una sostanza alcalina alle foglie di coca provoca un aumento fino a 10 volte della concentrazione di cocaina nel plasma sanguigno. La scultura, che è alta 19,7 centimetri, si trova nella collezione del Museum of the American Indian di New York. Viene attribuita alla prima epoca della cultura Moche, che fiorì lungo la costa settentrionale del Perù fra il 100 a.C. e il 100 d.C.

La cocaina è un alcaloide che fa parte di un vasto gruppo di sostanze vegetali che comprende anche la nicotina, la caffeina e la morfina. In natura la cocaina si trova in quantità significative soltanto nelle foglie di due specie dell'arbusto coca. *Erythroxylum coca* richiede un clima umido, tropicale, secondo Timothy Plowman del Field Museum of Natural History di Chicago, sarebbe originaria delle Ande peruviane, anche se ora cresce lungo tutte le regioni montagnose orientali delle Ande, in Ecuador, in Perù e in Bolivia. La concentrazione dell'alcaloide in queste foglie può arrivare fino all'1,8 per cento. *E. coca* fu coltivata dagli Incas ed è tuttora la principale fonte del commercio clandestino in atto su scala mondiale.

E. novogranatense, l'altra specie di *Erythroxylum* ricca di cocaina, viene coltivata nelle asciutte regioni montagnose della Colombia e lungo la costa caraibica dell'America Meridionale. La *truxillense,* o «Trujillo», varietà di questa seconda specie, è attualmente coltivata lungo le coste settentrionali del Perù e nell'asciutta vallata del fiume Marañón, un affluente del Rio delle Amazzoni, nel Perù nordorientale. Le sue foglie sono raccolte per essere esportate legalmente a Maywood, N. Y., dove la cocaina viene estratta per scopi farmaceutici legali, mentre il rimanente materiale della foglia viene convenientemente preparato per essere usato come aromatizzante della Coca-Cola.

Recenti ritrovamenti archeologici in Ecuador provano che l'esperienza umana della cocaina risale ad almeno 5000 anni addietro, molto prima cioè che venisse instaurato l'impero Inca. Per gli Incas la coca era una pianta di origine divina e un simbolo del loro rango politico e sociale. Secondo la mitologia, il dio Inti aveva creato la coca per alleviare la fame e la sete degli Incas, i quali si consideravano discendenti degli dei. Lo stato Inca controllava di fatto ogni aspetto della vita quotidiana, inclusi la coltivazione e l'uso della coca. La coca veniva masticata soprattutto dalle classi dominanti, anche se occasionalmente veniva distribuita ai soldati, agli operai e ai messaggeri. La masticazione non autorizzata delle foglie veniva considerata un sacrilegio.

Le prime notizie sulla masticazione della coca sono arrivate in Europa quasi in concomitanza con la scoperta del Nuovo Mondo. Le lettere di Amerigo Vespucci, pubblicate nel 1507, parlano della pratica degli indios di masticare le foglie di coca e di aggiungere ceneri al bolo. Il principale componente della cenere è il carbonato di calcio ($CaCO_3$), il quale aumenta gli effetti soggettivi e metabolici delle foglie di coca; fra gli indios delle Ande questa pratica è ancora in uso. Altri osservatori europei prima furono disorientati, poi provarono repulsione e infine restarono impressionati di fronte all'uso della coca da parte degli indigeni. Un manoscritto datato 1613 di Don Phelipe Guaman Poma de Ayala descrive la masticazione della coca come un'usanza sociale non autorizzata messa in atto dagli indios prima di cominciare a lavorare.

Altri resoconti di esploratori e cronisti sull'uso della masticazione della coca furono più entusiastici. Venne riferito che la coca provocava un sensazionale aumento della resistenza alla fatica e che permetteva ai soggetti di compiere un lavoro faticoso con poco cibo e ad altitudini notevoli. Gli imprenditori spagnoli in America Meridionale ebbero agio di constatare l'utilità della coca nel reclutamento di lavoratori indios e Filippo II di Spagna approvò ufficialmente questa loro pratica dichiarando che la coca era necessaria al benessere degli indios.

Se le foglie di coca conservassero la loro efficacia nonostante l'essiccamento subìto e il lungo viaggio di trasporto, come fanno le foglie di tabacco e i chicchi di caffè, il loro ruolo nella società europea e nordamericana sarebbe stato assai diverso da quello che oggi ricoprono. Campioni botanici di coca arrivarono in Europa poco dopo l'esplorazione spagnola, e foglie di coca continuarono a essere esportate in Europa in piccole quantità durante i secoli XVII, XVIII e all'inizio del XIX secolo. A dispetto delle entusiastiche descrizioni degli effetti della coca presso gli indios, gli europei non adottarono la pratica della sua masticazione. In termini di importanza culturale, la effettiva traversata dell'Atlantico da parte della coca ha dovuto attendere lo sviluppo di metodi di isolamento di composti chimici, avvenuto nel XIX secolo. Quando con questi metodi diventarono disponibili l'alcaloide puro e le sostanze affini, ciò avvenne in un contesto sociale assai diverso da quello in cui, durante il XVI secolo, avevano attecchito il tabacco e il caffè. La moderna storia sociale della cocaina è stata in gran parte determinata da modi di pensare che si erano venuti affermando alla fine del XIX secolo e dalle reazioni a tali modi di pensare.

Il commercio illegale di cocaina ha origine nelle Ande orientali dove le foglie di coca vengono raccolte *(aree in colore della cartina a sinistra).* Il raccolto, trasformato in «pasta» di cocaina, per vie tortuose raggiunge laboratori clandestini specialmente in Colombia *(frecce nella cartina a sinistra)* dove viene trasformata in cloridrato di cocaina. Le vie del contrabbando partono dalle città portuali della Colombia e dalle piste di volo delle città dell'interno verso vari porti degli Stati Uniti e d'Europa. Nella cartina a destra sono indicate con frecce in colore le vie con il maggior volume di traffico di cocaina; in nero sono indicate le vie secondarie e in grigio le rimanenti. Giunta ai paesi di distribuzione, la cocaina viene tagliata con vari adulteranti, in una proporzione che varia tra il 10 e l'85 per cento. Si ritiene che nel 1980 siano entrate negli Stati Uniti tra 40 e 48 tonnellate di cocaina, per un valore compreso tra 27 e 32 miliardi di dollari ai prezzi di vendita al dettaglio.

Il chimico tedesco Friedrich Gaedcke è stato probabilmente il primo ad isolare l'alcaloide cocaina dalla foglia di coca nel 1855, quando dal residuo di distillazione di estratto di coca isolò una sostanza oleosa e piccoli cristalli. Albert Niemann dell'Università di Göttingen fu il primo a caratterizzare chimicamente la sostanza. Niemann notò che la cocaina aveva un gusto amaro e che in breve tempo rendeva insensibile la lingua. Fin dal 1880 Vasili von Anrep, un nobile russo, medico all'Università di Würzburg, aveva notato di essere diventato insensibile alla puntura di uno spillo dopo essersi iniettato cocaina per via sottocutanea. Intanto, poiché sembrava che la cocaina avesse un'azione opposta a quella delle droghe che deprimono il sistema nervoso centrale, alcuni medici americani ne studiarono il possibile uso quale antidoto all'assuefazione alla morfina e all'alcool.

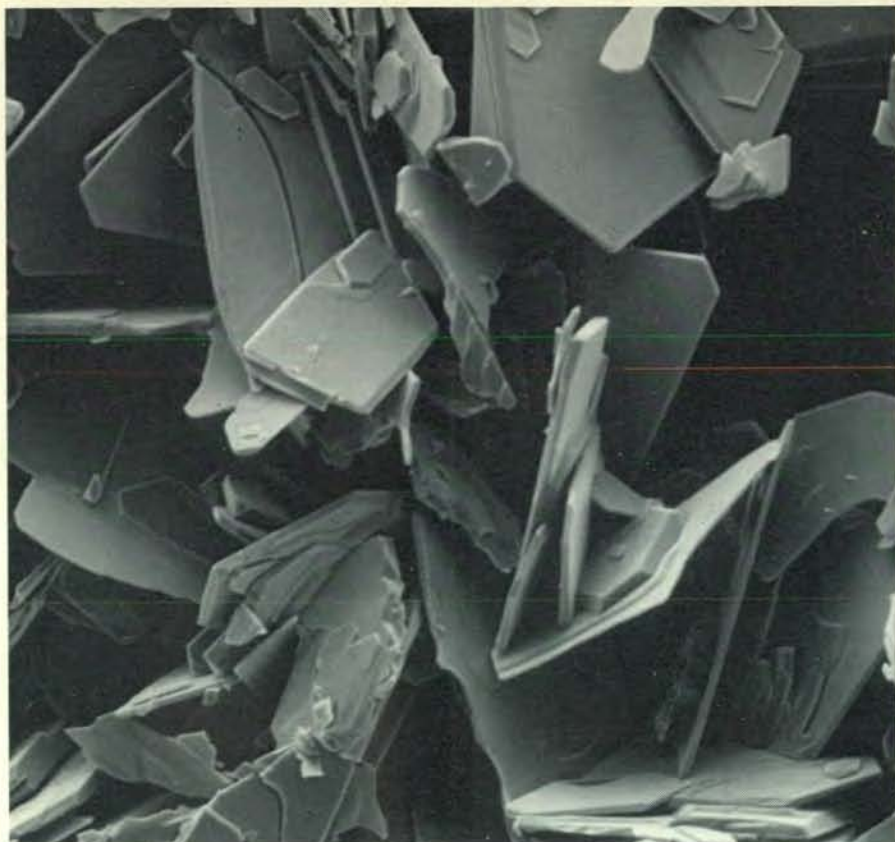
L'arbusto della coca *(Erithroxylum coca)* fiorisce nelle umide foreste tropicali dei contrafforti orientali delle Ande a una altitudine compresa tra i 500 e i 1000 metri. L'alcaloide cocaina si trova nelle foglie dell'arbusto in una concentrazione che può arrivare all'1,8 per cento. Nelle coltivazioni la pianta viene portata a una altezza di circa due metri in modo da facilitare la raccolta delle foglie. La cocaina si trova anche in altre specie del genere *Erithroxylum,* ma in minor quantità.

Nel 1884 le notizie entusiastiche sulla cocaina attirarono l'attenzione di Sigmund Freud, che era allora giovane assistente nella sezione neurologica dell'ospedale di Vienna. Nonostante von Anrep fosse uno dei suoi insegnanti, Freud concentrò il suo interesse sulla cocaina non come anestetico locale, ma piuttosto come stimolante del sistema nervoso centrale e come aiuto per vincere la dipendenza da morfina. Egli stesso ingerì cocaina e pubblicò, con il titolo di *Sulla cocaina* (Newton Compton, Roma, 1979), una rassegna sulla letteratura disponibile sulla coca e sulla cocaina.

Sulla cocaina suscitò un vivace interesse. La pubblicazione esprimeva un'opinione altamente favorevole nei riguardi della droga, e in lettere dello stesso periodo occasionalmente Freud scriveva: «una piccola dose mi tira su in modo meraviglioso». Nonostante egli avesse suggerito la possibilità dell'uso terapeutico di cocaina per attenuare il dolore provocato dall'infiammazione, toccò al suo amico e collega Carl Koller di riconoscere tutte le implicazioni delle proprietà della cocaina quale anestetico locale. Koller la usò nell'occhio in occasione di interventi chirurgici e in questo modo dimostrò inequivocabilmente la sua efficacia come anestetico locale nella pratica chirurgica.

Non altrettanto bene riuscì il tentativo di dimostrare che la cocaina e la morfina possono essere usati quali antagonisti e cioè per contrastare l'uno gli effetti dell'altro. Il chirurgo americano William Stewart Halsted fu una delle prime vittime della dipendenza da cocaina in conseguenza di esperimenti che egli aveva condotto su se stesso per dimostrare che la droga poteva agire come «neurobloccante». Il suo isolamento a bordo di una goletta non bastò per eliminarne l'assuefazione e Halsted fu trattato con morfina nella speranza che gli effetti farmacologici di questa, chiaramente opposti, potessero cancellare il suo desiderio di cocaina. Lo sfortunato risultato fu che Halsted aggiunse alla dipendenza psicologica dalla cocaina una dipendenza fisica dalla morfina che non lo lasciò più fino alle fine dei suoi giorni.

Quando l'inefficacia della cocaina nel trattamento della dipendenza da morfina fu chiara e cominciarono a circolare notizie gravemente preoccupanti sulla cocaina, l'entusiasmo di Freud per la droga cominciò a vacillare. Nonostante egli continuasse a godersi il proprio privato consumo della droga, restò tuttavia turbato quando, tre anni dopo le sue prime ricerche sulla cocaina, Albrecht Erlenmeyer lo accusò di avere scatenato «il terzo flagello dell'umanità» (dopo l'alcool e gli oppiati).

Le descrizioni degli effetti della cocaina sull'uomo fornite da Freud e dai suoi collaboratori restarono insuperate per circa 90 anni. La maggior parte dei lavori scientifici successivi sull'argomento furono rivolti a chiarire i meccanismi biochimici che potevano spiegare le osservazioni di Freud; circolarono molte dicerie, aneddoti e invettive. L'interesse per la cocaina derivava soprattutto dalle sue proprietà quale anestetico locale, e in conseguenza di questo interesse la cocaina fu uno dei primi alcaloidi a essere sintetizzato. Richard Willstätter e i suoi colleghi dell'Università di Monaco ottennero la sintesi nel 1923, ma la struttura spaziale della molecola di cocaina rimase sconosciuta fino al 1955, quando fu stabilita da E. Von Hardeggar e da Hans Ott.

La figura mostra cristalli di cloridrato di cocaina ingranditi 600 volte in una fotografia realizzata con il microscopio elettronico a scansione da Alan Pooley, del Peabody Museum della Yale University, e da uno degli autori (Byck). Raramente i consumatori di cocaina si imbattono nella sostanza effettivamente pura. I cristalli si disciolgono nelle mucose e vengono assorbiti nel sangue.

Un anestetico locale blocca la conduzione dell'impulso nervoso sensoriale quando viene applicato direttamente sul nervo. Mentre molte sostanze possono bloccare la conduzione nervosa in quanto ledono o distruggono definitivamente il tessuto nervoso, un anestetico locale è diverso delle altre sostanze in quanto produce un effetto temporaneo e reversibile. Molte delle droghe, dotate di proprietà anestetiche locali, sintetizzate dall'inizio di questo secolo, quali la procaina (Novocaina) e la lidocaina (Xylocaina), sono strutturalmente simili alla cocaina e agiscono con un meccanismo analogo.

Un anestetico locale inibisce la trasmissione nervosa in quanto altera la membrana della cellula nervosa. Quando la cellula è nel suo stato di riposo, c'è una differenza di potenziale di 60-70 millivolt tra l'interno dell'assone (il filamento che emerge dal corpo cellulare e conduce l'impulso) e il fluido che circonda l'assone stesso. L'assoplasma all'interno dell'assone è caricato negativamente rispetto al fluido che si trova all'esterno della membrana cellulare. Quando la cellula viene stimolata, il segnale conosciuto come potenziale d'azione percorre tutto l'assone come un'onda di depolarizzazione elettrica. Una lieve depolarizzazione della membrana dell'assone provoca, di solito, un notevole aumento transitorio della permeabilità della membrana stessa agli ioni sodio carichi positivamente che si trovano all'esterno della cellula. A mano a mano che questi ioni attraversano i canali della membrana ed entrano nell'assoplasma, la differenza di potenziale tra l'assoplasma e il fluido circostante si riduce. Questo movimento di cariche è un processo autolimitante e, poco dopo l'inizio del potenziale d'azione, potassio, calcio e altri ioni carichi positivamente fuoriescono dall'assone e ristabiliscono l'equilibrio elettrochimico della cellula.

Non si sa con precisione in che modo un anestetico locale modifichi la membrana dell'assone, ma oggi si pensa che le molecole dell'anestetico si dissolvano nella matrice lipidica della membrana stessa e si fissino su particolari siti recettori all'interno dei canali del sodio. Probabilmente la presenza di molecole di anestetico nei siti recettori interferisce con l'apertura dei canali stessi, per cui viene inibito il passaggio degli ioni sodio attraverso la membrana. In questo modo viene impedita la depolarizzazione della membrana e l'impulso nervoso viene bloccato. L'azione dell'anestetico termina quando le molecole dell'anestetico si spezzano e si diffondono nella circolazione sanguigna.

Un'altra proprietà della cocaina che è stata valorizzata nella pratica clinica è la sua tendenza a provocare vasocostrizione in caso di applicazione topica. La cocaina è l'unico anestetico locale che possiede questo effetto e fino a un'epoca recente è stato l'anestetico di scelta nella chirurgia oculare a causa della sua capacità di limitare la circolazione sanguigna. Si è constatato, peraltro, che questa riduzione della circolazione può danneggiare la superficie dell'occhio e perciò la cocaina non è stata più raccomandata in oculistica. Conserva invece un suo ruolo nella chirurgia delle mucose, per esempio dell'orecchio, del naso e della gola, e in procedimenti che richiedono l'introduzione di un tubo attraverso il naso o la gola.

La vasocostrizione sanguigna è una delle varie conseguenze a livello periferico della stimolazione del sistema nervoso simpatico. Le droghe che riproducono gli stessi effetti dei nervi simpatici sono definite droghe simpaticomimetiche. La cocaina è un tipico agente simpaticomimetico in quanto aumenta la frequenza cardiaca, innalza la pressione sanguigna, aumenta la temperatura corporea e provoca la dilatazione delle pupille.

La trasmissione continua dei segnali lungo le vie nervose simpatiche è essenziale per la vita. Nell'assone l'eccitamento assume la forma di un potenziale di azione, ma quando il potenziale d'azione raggiunge la parte terminale dell'assone il segnale viene inviato alla cellula vicina per mezzo di un meccanismo diverso e cioè mediante la liberazione di un neurotrasmettitore quale la dopammina o la noradrenalina. Il neurotrasmettitore viene immagazzinato in vescicole all'interno dell'assone in vicinanza della sinapsi (zona di giunzione tra l'assone di una cellula nervosa e i dendriti di quella successiva). Quando l'assone conduce lo stimolo, le vescicole si fondono con la membrana cellulare liberando il neurotrasmettitore nella fessura sinaptica tra le cellule. Le molecole liberate stimolano il neurone immediatamente successivo, attaccandosi a siti recettori che sono situati sui suoi dendriti.

Normalmente le molecole di neurotrasmettitore che non stimolano la cellula nervosa immediatamente successiva vengono decomposte da enzimi, diffondono nei tessuti adiacenti oppure vengono pompate nuovamente all'indietro, all'interno della cellula terminale che le ha liberate. Questa azione di pompa viene denominata riassunzione *(reuptake)*. Quando molecole di cocaina sono presenti nella fessura sinaptica, tuttavia, esse inibiscono il meccanismo di riassunzione e in tal modo producono gli stessi effetti che avrebbe un'ulteriore liberazione di molecole di neurotrasmettitore. Il neurotrasmettitore tende a rimanere nella fessura, dove può continuare a stimolare i recettori che si trovano nei dendriti della cellula nervosa immediatamente successiva.

In passato si riteneva che gli effetti di eccitazione psicologica della cocaina dovessero attribuirsi esclusivamente alla sua capacità di bloccare la riassunzione di noradrenalina. Oggi si sa che non è così. Molte delle droghe antidepressive usate in psichiatria, come ad esempio l'amitriptilina (Elavil), bloccano la riassunzione

dei neurotrasmettitori, ma non sono eccitanti e non producono euforia.

D'altra parte non ci sono sufficienti ragioni per pensare che un anestetico locale come la lidocaina possa provocare euforia e tuttavia le nostre ricerche hanno dimostrato che la provoca. Una dose singola di lidocaina somministrata per via nasale non può essere distinta rispetto ad una stessa quantità di cocaina. Inoltre, altri gruppi hanno dimostrato su animali da esperimento che essi hanno le stesse reazioni nei confronti di iniezioni di procaina o di cocaina. Dal momento che le molecole di cocaina, lidocaina e procaina hanno una struttura simile, queste constatazioni possono suggerire semplicemente che droghe simili producono esperienze simili. Inoltre fanno pensare, d'altra parte, che certi tipi di euforia provocati dalla droga rappresentino interpretazioni generalizzate, da parte dei tossicomani, di un ampio spettro di sensazioni non comuni provocate dall'interazione di molteplici fattori psicologici e ambientali.

La struttura della cocaina assomiglia a quella degli anestetici locali lidocaina (Xylocaina) e procaina (Novocaina). Tutte e tre le molecole contengono un gruppo aromatico che le rende solubili nei tessuti ricchi di sostanze adipose, come ad esempio il cervello, e un gruppo amminico, con un atomo di azoto. La lidocaina e la procaina sono tra gli adulteranti con cui viene talora tagliata la cocaina venduta dagli spacciatori. Gli autori hanno constatato che la cocaina, somministrata per via nasale, può non essere distinta dalla lidocaina anche da un esperto consumatore.

Gli effetti di una droga dipendono dalla dose, dalla forma in cui viene assunta e dalla via di introduzione. Molti effetti variano anche a seconda della frequenza con cui la droga viene assunta: il corpo può accumulare la droga oppure sviluppare una tolleranza nei suoi confronti. Gli effetti delle droghe psicoattive variano altresì in rapporto con le aspettative di chi le assume, con l'ambiente e con le circostanze in cui sono assunte, infine con la storia e la personalità del fruitore. L'intensità di queste ultime tre variabili nel determinare gli effetti complessivi di una droga rende difficile una stima corretta di quanta parte dell'esperienza della cocaina sia realmente da attribuire alla sostanza stessa e di quanta parte, invece, debba essere attribuita all'ambiente o alle aspettative.

Quando la cocaina viene assunta a «uso voluttuario» l'aspettativa e la mitologia che circondano la droga possono generare impressioni soggettive che sopprimono gli effetti non concordi con le aspettative. Inoltre è probabile che la cocaina acquistata sulla strada sia adulterata, ciò che complica la farmacologia della droga. L'adulterante può essere uno stimolante di basso costo, come l'amfetamina, un semplice carboidrato come il mannitolo o il lattosio, un anestetico come la procaina, e praticamente qualsiasi altra sostanza che l'acquirente non può immediatamente individuare. In altre parole quando la cocaina viene somministrata in condizioni sperimentali controllate, gli stimoli che possono intensificare o attenuare i suoi effetti sono solitamente assenti. La conoscenza degli effetti acuti della cocaina deriva necessariamente da un miscuglio di descrizioni individuali, che possono essere inficiate da pregiudizi o fantasticherie, e di dati sperimentali, i quali possono depurare l'esperienza della droga fino a renderla irriconoscibile.

Nemmeno in laboratorio le aspettative della persona cui sarà somministrata la droga possono essere eliminate totalmente. L'etica della sperimentazione delle droghe su soggetti umani ammette che nelle ricerche sui loro effetti siano impiegati soltanto esperti consumatori di cocaina. Ciò può alterare i dati sperimentali in quanto costoro hanno già maturato le loro aspettative riguardo alla cocaina in una situazione ambientale assai diversa da quella del laboratorio. Per ridurre ai minimi termini l'intervento di questi fattori di distorsione è importante che né lo sperimentatore né il soggetto conoscano la reale entità della dose somministrata (quando la sostanza viene realmente somministrata); inoltre, per verificare le aspettative, un placebo deve essere inframezzato alle dosi attive di cocaina.

Per dare un impianto organico all'inquadramento delle risposte dei nostri soggetti alla cocaina, abbiamo deciso di

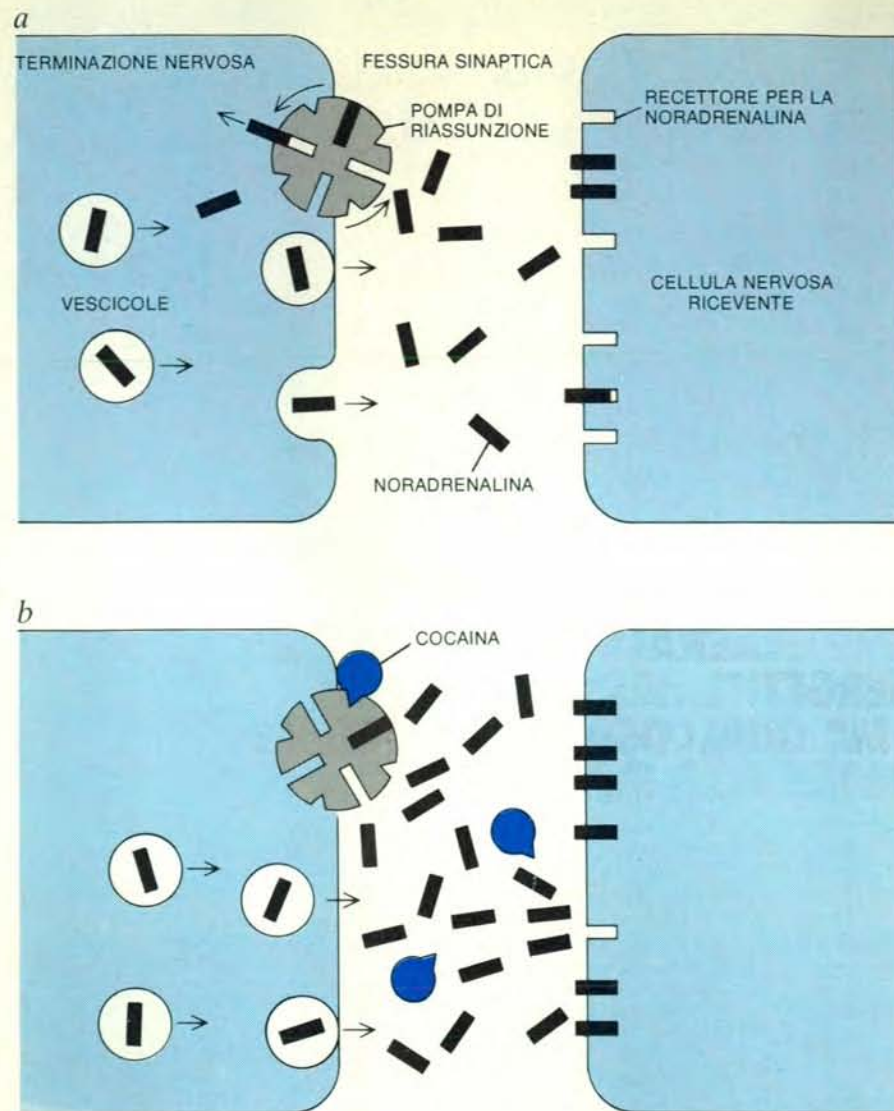
Gli anestetici locali inibiscono la trasmissione degli impulsi di un nervo sensoriale in quanto modificano la permeabilità della membrana dell'assone agli ioni sodio. Normalmente l'impulso si propaga come un'onda di depolarizzazione che passa lungo la membrana dell'assone; nel corso della depolarizzazione viene momentaneamente modificata la differenza di potenziale elettrico esistente tra l'assoplasma e il fluido circostante. Come illustra il diagramma schematico *(a)*, un canale di dispersione consente agli ioni carichi positivamente che si trovano sia all'interno sia all'esterno dell'assone di attraversare la membrana dell'assone in ambedue le direzioni. Tuttavia, la concentrazione degli ioni sodio *(cerchi in colore)* si mantiene più alta all'esterno dell'assone rispetto al suo interno a causa della presenza di una pompa del sodio (non illustrata nel diagramma), mentre, al contrario, la concentrazione del potassio, del calcio e degli altri ioni *(cerchi in grigio)* risulta più alta all'interno dell'assone. (Nel diagramma le varie concentrazioni sono state esagerate per maggiore chiarezza.) Allo stato di riposo, l'equilibrio viene raggiunto quando l'interno dell'assone è carico negativamente rispetto all'esterno. Quando il nervo viene stimolato *(b)*, cominciano ad aprirsi delle «porte di attivazione» che si trovano nei canali del sodio della membrana e che permettono soltanto ai piccoli ioni sodio di penetrare dentro l'assoplasma. In conseguenza di ciò la differenza di potenziale tra l'interno e l'esterno dell'assone si riduce. Quando i canali del sodio vengono chiusi da «porte di inattivazione», altri ioni fuoriescono dall'assone per ripolarizzare la membrana. In presenza di cocaina *(c)* gli ioni sodio non possono passare attraverso il canale del sodio, innanzitutto perché esso è bloccato da parte della cocaina, ma anche perché la porta di inattivazione rimane chiusa più saldamente. Come conseguenza di tale meccanismo l'impulso nervoso resta bloccato.

ricorrere alla misura della cocaina effettivamente presente nel corpo. Nel 1975 David Bailey e Peter Jatlow della Yale School of Medicine hanno sviluppato una metodica sensibile e specifica per la determinazione della concentrazione di cocaina nel plasma sanguigno. Il loro metodo si basa sull'impiego di un gascromatografo, uno strumento in cui le molecole di un gas semplice vengono forzate attraverso una colonna impaccata con un materiale adsorbente. Se nel gas ci sono tipi vari di molecole esse vengono adsorbite e liberate con diversa velocità, cosicché una sostanza può essere identificata dal tempo caratteristico impiegato per attraversare la colonna. Per aumentare la sensibilità del dispositivo alla cocaina, Bailey e Jatlow hanno installato un rivelatore particolarmente sensibile ai composti contenenti azoto, come appunto la cocaina. Questo metodo ha consentito per la prima volta di misurare la concentrazione della cocaina nel sangue anche per valori dell'ordine di cinque nanogrammi per millilitro. In tal modo è possibile seguire gli effetti delle dosi usualmente adoperate dai cocainomani e prima degli interventi chirurgici.

Fino al momento del lavoro di Bailey e Jatlow la determinazione della concentrazione di cocaina nel sangue era stata ostacolata dalla presenza nel sangue stesso della pseudocolinesterasi, un enzima che è responsabile della demolizione del neurotrasmettitore acetilcolina, così come della cocaina. Aggiungendo ioni fluoruro al plasma fresco Bailey e Jatlow hanno scoperto che l'azione dell'enzima può essere inibita. Il progresso tecnico ha consentito a sua volta la determinazione dei livelli plasmatici di cocaina presso i masticatori di coca dell'America Meridionale, in condizioni ambientali che non sono direttamente accessibili agli strumenti di misura dei laboratori.

Sotto forma di cloridrato la cocaina viene rapidamente assorbita attraverso le mucose ed entra in circolo. Per motivi sconosciuti la via di introduzione più popolare è quella attraverso la mucosa nasale, anche se il nostro gruppo ha dimostrato che i cristalli di cloridrato di cocaina si sciolgono e sono assorbiti altrettanto bene nel tratto gastrointestinale. La cocaina può essere introdotta anche per via endovenosa, oppure può essere assorbita attraverso i polmoni fumando la sostanza alcaloide che i consumatori chiamano *free-base*. Attualmente nell'America Meridionale i giovani dei centri urbani fumano un prodotto intermedio della preparazione del cloridrato di cocaina chiamato pasta di cocaina. La pasta di cocaina è costituita da cocaina *free-base* in una percentuale tra il 30 e il 90 per cento. Negli Stati Uniti sono in vendita kit di laboratorio per dilettanti per convertire il cloridrato di cocaina in cocaina *free-base,*

a

b

L'azione simpaticomimetica della cocaina è conseguente al blocco della riassunzione *(reuptake)* di neurotrasmettitori, come ad esempio la noradrenalina, a livello delle sinapsi del sistema nervoso simpatico. (Il sistema nervoso simpatico controlla funzioni quali la frequenza cardiaca e la pressione arteriosa.) Quando a livello della terminazione nervosa *(a)* le molecole di neurotrasmettitore vengono liberate da parte delle vescicole, esse attraversano la fessura sinaptica e stimolano la cellula nervosa successiva. Di regola parte delle molecole di neurotrasmettitore che si trovano nella fessura sinaptica vengono ripompate verso il nervo che le ha liberate. In presenza di cocaina l'azione della pompa di riassunzione viene bloccata *(b)* e l'azione stimolatrice delle molecole di neurotrasmettitore aumenta con l'aumentare della concentrazione nella fessura sinaptica.

che è meno suscettibile all'alterazione termica. La cocaina *free-base* evapora sotto l'azione di gas caldi generati da un *base pipe* (in pratica, un cannello) e in questo modo vengono inalati i vapori.

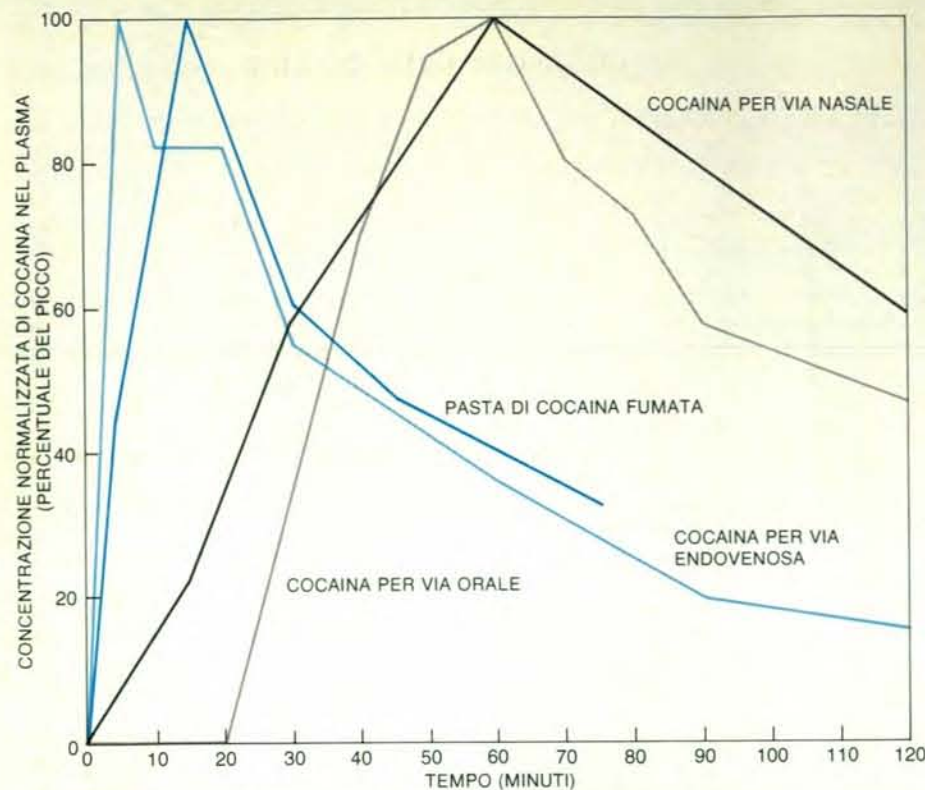
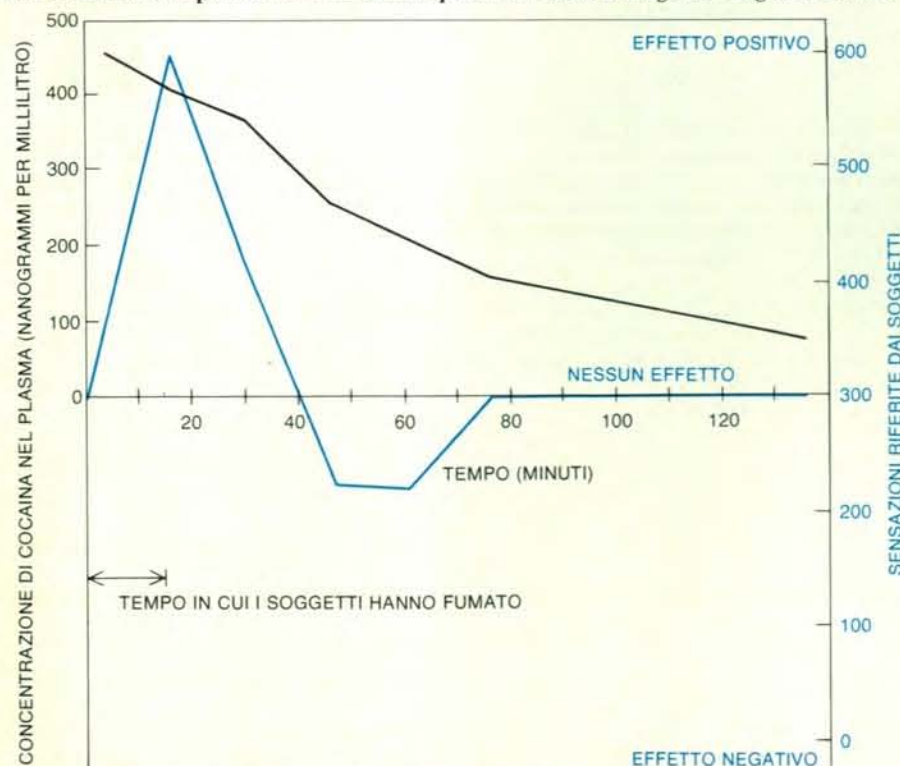
In alcune esperienze di laboratorio la cocaina è stata somministrata per via endovenosa in quanto in questo modo può essere stabilito con precisione il momento in cui la droga viene completamente assorbita. La concentrazione di cocaina nel sangue raggiunge il suo massimo immediatamente dopo l'iniezione e successivamente diminuisce con una vita media tra i 45 e i 90 minuti. I soggetti in esperimento dichiarano di provare un'intensa sensazione di euforia seguita talvolta da una sensazione di «crollo», cioè di intensa depressione, e dal desiderio smodato di una maggiore quantità di cocaina.

Gli effetti della cocaina per via endovenosa sono talmente brevi che non è possibile realizzare la maggior parte dei test sui suoi effetti, test che richiedono un certo tempo. Noi abbiamo perciò preferito studiare gli effetti fisiologici e psicologici della cocaina introdotta per via nasale. Non abbiamo richiesto ai soggetti di aspirare i cristalli di cocaina con una cannuccia di paglia o con una banconota da un dollaro arrotolata; né abbiamo seguito la comune pratica della strada di somministrare una quantità relativamente piccola di cocaina (circa 30 milligrammi) a intervalli regolari. Al contrario, abbiamo dato, con uno spray nasale, una dose unica, ma sufficientemente alta per poter esaminare minuziosamente l'andamento dell'azione della droga.

Gli effetti fisiologici della cocaina introdotta in questo modo non sono violenti. Non ci ha sorpreso la constatazione che dosi più alte producevano livelli plasmatici più alti, anche se, in apparente contraddizione con la reputazione della cocaina quale droga a vita breve, il picco di concentrazione plasmatico veniva raggiunto circa un'ora dopo la somministrazione della droga. Dosi tra 25 e 100 milligrammi aumentano la frequenza cardiaca di 8-10 battiti al minuto e innalzano la pressione sanguigna, sia sistolica sia diastolica, di circa 10 millimetri di mercurio. Questi effetti raggiungono il loro massimo intorno a 15-30 minuti dal momento della somministrazione della droga. Alla dose relativamente alta di due milligrammi per chilogrammo di peso corporeo (cioè circa 140 milligrammi in una persona di peso medio) si registra un lieve aumento della temperatura, ma non necessariamente dilatazione delle pupille.

Sebbene la cocaina ad alte dosi, come tutti gli altri anestetici locali, produca convulsioni, non si registra un apprezzabile effetto sull'elettroencefalogramma alla dose di due milligrammi per chilogrammo di peso corporeo. Peraltro, negli animali la somministrazione ripetuta di piccole dosi di cocaina può, ad un certo punto, provocare convulsioni. L'uso cronico di cocaina può provocare la brusca insorgenza di un comportamento apparentemente psicopatico. Alcuni ricercatori hanno concluso che ciò avviene per un «effetto somma», per cui una serie di piccole dosi produce un tipo di risposta quale normalmente ci si aspetta da una dose molto più grande.

Uno dei principali motivi del nostro interesse per la cocaina è la sua capacità di provocare in modo attendibile precise sensazioni nella maggior parte della gente. Per definire gli effetti psicologici della cocaina abbiamo adottato vari test comunemente impiegati nello studio dell'umore e del dolore. Nel corso di uno di tali test il soggetto viene invitato a indicare un punto su una linea contrassegnata in modo tale da rappresentare sei livelli di sensazione di euforia o di depressione. Nel corso di un altro test viene data al soggetto una lista di aggettivi, quali «vivace», «diffidente», «inquieto», «attento», «sonnolento» e «scoraggiato», insieme con una scala graduata nella quale egli dovrà indicare il grado in cui ciascuno di tali aggettivi corrisponde alle sensazioni che sta provando.

In esperimenti condotti in altri laboratori i soggetti non erano riusciti a distinguere una piccola dose di cocaina (10 milligrammi) da un placebo. Invece quando una dose tra 25 e 100 milligrammi veniva assunta per via nasale tutti i soggetti riferivano, in modo attendibile, un picco di euforia nel giro di 15-30 minuti. Alla distanza di 45-60 minuti dalla dose di 100 milligrammi, due soggetti su sette provavano ansietà, depressione, stanchezza e desiderio di altra cocaina. Si registra spesso la caduta precipitosa in una condizione di estremo sconforto dopo che la cocaina è stata fumata o assunta per via endovenosa, ma ciò avviene raramente dopo sommini-

La figura illustra l'andamento della concentrazione di cocaina nel plasma sanguigno per dosi di cocaina introdotte nel corpo attraverso quattro vie. Le concentrazioni sono state normalizzate e tutti i valori sono espressi in percentuale del valore di picco. Una volta raggiunta la concentrazione di picco, la cocaina ha sulla circolazione sanguigna una vita media di circa un'ora. Probabilmente differenze nella velocità con cui la concentrazione plasmatica di cocaina raggiunge il proprio picco a seconda della via attraverso la quale viene introdotta sono causa di differenti effetti psicologici della cocaina stessa. Quando la cocaina viene iniettata o fumata si ha un più rapido aumento della sua concentrazione plasmatica in confronto a quanto avviene se la droga viene ingerita o inalata.

In condizioni sperimentali si è visto che sensazioni di depressione succedono all'euforia iniziale in soggetti che hanno fumato pasta di cocaina, e che ciò avviene mentre la concentrazione plasmatica di cocaina è ancora relativamente elevata. La depressione è caratterizzata da ansia, angoscia, stanchezza e desiderio di ulteriori dosi di cocaina. Probabilmente in questi soggetti le sensazioni di depressione hanno raggiunto la propria acme un poco prima del momento di massima concentrazione plasmatica, ma i primi test psicologici sono stati effettuati dopo le analisi plasmatiche. Perciò la sequenza temporale dei rispettivi picchi non può essere considerata un'acquisizione definitiva.

strazione per via nasale. Per lo più l'uso delle droghe viene associato dalla gente con un solo specifico effetto, ma sia dalla esperienza della strada sia da quella dei laboratori risulta chiaramente che gli effetti piacevoli della cocaina a dosi relativamente piccole sono sostituiti da sensazioni spiacevoli nel caso di dosi maggiori.

Quando l'andamento degli effetti psicologici della cocaina viene paragonato con l'andamento della sua concentrazione nel plasma sanguigno, si constata che lo stato di euforia raggiunge il suo massimo pochi istanti prima che la concentrazione plasmatica cominci a scendere. Gli effetti euforici scompaiono molte ore prima che la concentrazione plasmatica ritorni a zero; una tale correlazione temporale tra le sensazioni soggettive e la concentrazione plasmatica può essere messa in evidenza anche con altre droghe, quali l'alcool o le benzodiazepine. Quando la cocaina viene fumata o iniettata in vena, la depressione può comparire quando la concentrazione plasmatica è ancora la metà della concentrazione massima.

La mancanza di coincidenza tra l'andamento della concentrazione plasmatica e l'andamento della sensazione di euforia fa pensare che gli effetti psicologici della cocaina siano correlati alla velocità con cui varia la sua concentrazione nel plasma, piuttosto che al livello assoluto di tale concentrazione. Poiché la concentrazione plasmatica varia molto più velocemente quando la cocaina viene fumata o iniettata, in confronto a quanto avviene quando viene aspirata o assunta per via orale, se la velocità di variazione è veramente il fattore che controlla il risultato dovremmo attenderci effetti psicologici molto più potenti quando la cocaina è stata fumata o iniettata. In effetti le sensazioni riferite dai soggetti in esperimento concordano con questa ipotesi. Si può anche pensare che la rapida diminuzione delle sensazioni di euforia nonostante l'alta concentrazione plasmatica di cocaina sia dovuta a una diminuita capacità di risposta dei siti recettori, una volta che sono stati occupati, per cui si instaura una tolleranza acuta alla droga.

Un'altra ragione del fatto che le sensazioni di euforia aumentano e diminuiscono con tanta rapidità, mentre permane alta la concentrazione plasmatica di cocaina, potrebbe essere dovuta al fatto che la cocaina è facilmente solubile nel tessuto adiposo. A causa di questa solubilità la cocaina può passare rapidamente dalla circolazione sanguigna nel cervello; vi sono alcune prove che la concentrazione di cocaina è più alta nel cervello che nel plasma. Se la grande concentrazione iniziale di cocaina nel plasma genera un'alta concentrazione nel cervello, il susseguente stato di equilibrio sarà raggiunto se la cocaina diffonderà nuovamente dal cervello verso il plasma. In questo modo la concentrazione plasmatica rimarrà alta anche dopo che la cocaina ha cominciato a lasciare il cervello. Dunque ciò che si sa circa la correlazione tra concentrazione plasmatica ed euforia è, come minimo, coerente con l'ipotesi che l'euforia provocata dalla cocaina sia strettamente correlata con la concentrazione di questa nel cervello.

Nella valutazione degli effetti di una droga bisogna distinguere le conseguenze del suo uso sul comportamento individuale a lungo termine e i distinti fenomeni che possono essere osservati dopo una dose singola. L'uso cronico di una droga può avere conseguenze sia sociali sia farmacologiche. Se una persona viene totalmente coinvolta nella ricerca e nel consumo di una droga, la sua intera struttura esistenziale può cambiare. Quando la cocaina viene assunta regolarmente può provocare insonnia e perdita di appetito; in seguito a dosi elevate o al consumo cronico chi la usa può andare incontro a stati paranoidi di ansia. Se la cocaina provochi o meno una vera psicosi è questione tuttora dibattuta, ma vi sono prove crescenti che possono verificarsi allucinazioni e paranoia se la droga viene assunta frequentemente a dosi elevate.

La cocaina produce dipendenza? Il modo in cui la domanda è posta sembra richiedere una risposta semplice, ma un sì o un no non renderebbero giustizia al groviglio di definizioni in termini medici, buon senso popolare, classificazione legale e recriminazione sociale che sottostà all'espressione «produce dipendenza». Una definizione medica di questa espressione, accettata da lungo tempo, deriva dalla descrizione degli effetti degli oppiati. Perché una droga sia classificata come capace di creare dipendenza è necessario che una persona sviluppi tolleranza nei suoi confronti e cioè uno stato per cui l'assunzione della medesima dose produca una risposta minore. Inoltre è necessario che la droga produca dipendenza fisica e cioè uno stato per cui è necessario ripetere la dose per evitare l'insorgenza di una sindrome da astinenza.

Nei termini di questa definizione la cocaina non è una sostanza che produce dipendenza. I consumatori di cocaina possono assumere la stessa dose tutti i giorni e ottenere lo stesso effetto. Nell'elettroencefalogramma e nel ritmo del sonno si possono mettere in luce segni di astinenza, ma essi non sono assolutamente drammatici se paragonati alle sindromi da astinenza associate alle tossicomanie da oppiati, barbiturici o alcool.

D'altra parte, la cocaina è certamente da classificare tra le sostanze che producono dipendenza nel senso più esteso del termine che oggi è stato adottato da molti farmacologi. E cioè la cocaina, a rigore, è una sostanza che dà assuefazione. In questo contesto bisogna distinguere il consumo di cocaina per via nasale dal suo uso sistematico per iniezione o mediante fumo di *free-base*. Nel consumo della cocaina per via nasale la dipendenza si manifesta in termini di consumo continuativo finché c'è a disposizione una scorta della sostanza e nella semplice astensione dal consumo quando questa scorta viene a mancare. Il modello di comportamento è paragonabile a quello che corrisponde all'esperienza di tanta gente con le noccioline e con le patatine fritte. Può interferire con altre attività dell'individuo, ma può anche rappresentare una fonte di piacere.

Viceversa il fumo o l'iniezione di cocaina può dar luogo a un comportamento di consumo e di ricerca continui della droga, che distrugge la personalità e ogni capacità produttiva. Inoltre il fumo della *free-base* è probabilmente nocivo anche all'apparato respiratorio in quanto provoca una vasocostrizione polmonare. Indubbiamente l'alto prezzo della cocaina al minuto, insieme con la relativamente grande quantità di sostanza che è necessaria per le iniezioni o per il fumo della *free-base*, sono fattori che limitano i danni che la gente altrimenti potrebbe autoprocurarsi.

Un consistente apparato burocratico deriva la propria esistenza dalla credenza popolare che la cocaina sia inequivocabilmente pericolosa. La Drug Enforce-

La diffusione della cocaina

Gli Stati Uniti d'America non sono soltanto il principale mercato mondiale del consumo di cocaina, ma anche il paese in cui la «neve» (usata allo stato puro ovvero in micidiali miscugli di più droghe) sta diventando il simbolo di uno stile di vita e di un comportamento di massa, che coinvolgono «quasi tutto il giovane mondo dello spettacolo e tramite la televisione si spandono attraverso l'America» (Furio Colombo). Nel corso di programmi seguiti da decine di milioni di spettatori, gli attori più popolari iniziano la gente alla «guerra bianca», illustrando le loro «comiche» avventure con la coca. Naturalmente queste avventure non hanno sempre una conclusione esilarante: Richard Pryor è recentemente rimasto ustionato e rintronato dallo scoppio della sua *base-pipe*, e John Belushi, sotto la spinta di una miscela di cocaina ed eroina, ha raggiunto le celesti praterie con la velocità di un siluro; ma nonostante queste esperienze negative, i consumatori di coca assumono ogni giorno di più il ruolo di «eroi casalinghi», mentre riscuote un grande successo di vendite il *Galateo della coca*, un prontuario che risponde a domande come: «Come si serve la cocaina agli ospiti?», «Quando si serve? Prima del pasto o dopo?»

È «una specie di bomba ad orologeria inserita dentro la vita e il costume d'America», ha commentato Norman Lear, un noto produttore televisivo.

L'Italia è stato uno dei primi paesi dove gli effetti della coca sono stati conosciuti e sperimentati.

Tra il 1850 e il 1860 (dunque prima di Freud), un medico famoso, Paolo Mantegazza, dopo aver provato personalmente l'infuso di foglie, comunicò eccitato ai suoi colleghi che avrebbe preferito «dieci anni con la coca ad un milione di secoli senza la coca». Peraltro nel nostro paese, complice il basso reddito capitario, il consumo di coca non è mai stato elevato.

	1977	1978	1979	1980	1981
MORFINA	0,5	2	76	268	82
EROINA	63	80	85	197	139
COCAINA	16	9	17	53	64
MARIJUANA	697	774	777	853	615
HASHISH	2117	3880	3453	4125	10 587

Quantitativi in chilogrammi di sostanze stupefacenti sequestrate in Italia negli ultimi anni. I dati riportati in questa tabella sono stati forniti dalla Guardia di finanza.

Un primato tuttavia l'Italia ce l'ha: l'Italia - e per l'Italia Milano - è la capitale europea del traffico internazionale di cocaina.

Nel 1977 la Guardia di finanza ne ha sequestrato 16 chili, 17 nel 1979, 53 nel 1980, 64 nel 1981. Nei primi tre mesi di quest'anno i chili sequestrati sono già 34. La coca parte dalla Bolivia, passa attraverso il Brasile e raggiunge la Svizzera, da dove, a bordo di treni, via lago oppure tramite spalloni, arriva a Milano, dove viene tagliata e poi dirottata a rifornire il mercato europeo.

Quanta se ne ferma in Italia?

Una stima corretta non è possibile. Ma è bene non farsi illusioni. Nel 1972, Amedeo Limentani, presidente dell'associazione psicanalitica inglese e vicepresidente dell'Associazione internazionale di psichiatria, a chi gli domandava quali fossero le sue previsioni riguardo alla possibile diffusione dell'eroina nel nostro paese, allora soltanto marginalmente colpito dalla piaga, rispondeva: «Roma entro dieci anni avrà la punta più alta, forse la punta più alta d'Europa, per un semplice motivo: Roma non esiste, è pietrificata nel tempo... A Roma non succede niente, ed essendo stupenda, bella, affascinante, ma colonialmente levantina, non può che assorbire acriticamente gli stimoli che vengono dall'esterno...».

Per l'eroina la profezia di Limentani si è avverata.

Ce ne ricorderemo ora che avanza la «moda» della cocaina? *(Lucio Rosaia)*

ment Administration statunitense impiega 1950 agenti a livello mondiale, con un budged operativo che durante l'anno fiscale conclusosi nel settembre 1981 ha toccato i 227 milioni di dollari. In tale periodo, gli agenti hanno dedicato a operazioni legate alla cocaina il 31 per cento delle ore-uomo della loro attività investigativa. Operatori dell'ente sono in attività in nove paesi dell'America Meridionale e in tutti questi paesi, esclusa la Colombia (dove anche il traffico di marijuana è considerato un problema), il principale oggetto della loro attività è la cocaina.

Per di più il Dipartimento di stato ha esercitato una grande quantità di pressioni diplomatiche per convincere i governi di questi paesi ad adottare linee di condotta gradite agli Stati Uniti, relativamente alla coltivazione della coca e alla severità delle leggi riguardanti il mercato della cocaina. Recentemente gli Stati Uniti hanno firmato trattati che impegnano il Perù e la Colombia all'eradicazione della coca; un accordo analogo non è stato raggiunto con la Bolivia. Presso le Nazioni Unite accordi per limitare l'incremento della produzione di coca in Perù e in Bolivia sono stati legati a incentivi economici e accompagnati da programmi di sviluppo di colture sostitutive.

Non è chiaro se la capacità potenziale della cocaina di dar luogo ad abuso giustifichi l'entità di questi sforzi. Morti per consumo di cocaina a scopo «voluttuario» sono rare. Nelle società delle Ande sono di comune riscontro livelli plasmatici di cocaina paragonabili a quelli dei consumatori di cocaina per via nasale e finora ci sono scarse prove di danni fisici connessi alla masticazione della coca. Presso queste culture l'uso religioso della cocaina non può essere classificato tra gli abusi di droga e lo zelo degli sforzi compiuti per l'eradicazione può privare gli indios di una sostanza tradizionale. La principale minaccia per la società europea e nordamericana è rappresentata dalla perdita di potenziale umano che potrebbe derivare da un consumo su larga scala. La condotta politica del problema della cocaina e i regolamenti tengono in scarso conto questi fatti e nel passato la legislazione sulla droga non è stata fondata né sulla conoscenza scientifica né sul buon senso. Abbiamo legalizzato sia l'alcool sia il tabacco, i quali, per quanto graditi a talune persone, costituiscono chiaramente un pericolo per la salute personale e collettiva.

La storia insegna che la cocaina è una droga che può essere facilmente consumata in eccesso. L'attuale mercato della cocaina ci dice che se la domanda è abbastanza grande può svilupparsi un'industria per soddisfarla. Dal punto di vista medico la cocaina è una droga relativamente innocua. Ma nelle mani di gente disinformata può portare a un comportamento suicida. Molte delle domande che il problema della cocaina solleva possono trovare risposta nella ricerca scientifica, ma le decisioni finali riguardo alla cocaina saranno basate su considerazioni di natura economica e politica, e non scientifica.

TEMI METAMAGICI

di Douglas R. Hofstadter

«National Enquirer» e «The Skeptical Inquirer»

«Ricercatori ed educatori perplessi rivelano... RAGAZZO RIESCE A VEDERE CON LE ORECCHIE».

«Incrocio tra esseri umani e piante... SCIENZIATI IN PROCINTO DI CREARE UOMINI VEGETALI... Le bizzarre creature potrebbero fare tutto quello che si vuole».

«L'ipnosi rivela: alieno venuto dallo spazio condivide la mente e il corpo di una donna».

— Titoli dal «National Enquirer»

Da ragazzi vi siete mai chiesti perché nei fumetti le frasi dichiarative termmavano sempre con punti esclamativi? Davvero tutte quelle affermazioni erano sorprendenti? Davvero i personaggi erano così elettrizzanti? Naturalmente no. Quei punti esclamativi erano un trucco psicologico per dare più brio alla storia.

Il «National Enquirer», uno dei giornali più scandalistici del paese, usa un trucco analogo. Ogni volta che stampa un articolo che strombazza la scoperta di qualche fenomeno bizzarro e inaudito, invece di terminarlo con un punto esclamativo lo termina (o lo inizia) con un riferimento a «ricercatori perplessi», «scienziati sconcertati» o altri sapienti analogamente stupiti. È un ornamento messo lì perché la storia sembri più credibile.

Ma è davvero così? Cosa vogliono in realtà gli editori? Che la storia appaia credibile o che la storia appaia incredibile? Si direbbe che vogliono entrambe le cose: vogliono che la storia sembri il più peregrina possibile e vogliono che abbia un'apparenza di autenticità. Il loro titolo ideale dovrebbe contenere una contraddizione: impossibilità accoppiata a certezza. In breve, il nonsenso confermato.

Cosa si può dire di titoli come quelli riportati in precedenza? O del fatto che questa pubblicazione si vende a milioni ogni settimana e che la gente ingoia le sue storie con la stessa voracità con cui mangia patatine? O del fatto che quando è esaurita hanno a disposizione altri prodotti dello stesso livello culturale come «National Examiner», «Star», «Globe», «Weekly World News». Cosa ne pensate?

La vostra prima reazione è probabilmente quella di sogghignare e lasciar perdere quelle storie in quanto stupide. Ma come fate a sapere che sono stupide? Pensate che anche *questa* sia una domanda stupida? Cosa pensate degli articoli stampati in questa rivista? Vi fidate? Qual è la differenza? È semplicemente una differenza nello stile di pubblicazione? È sufficiente il formato tabloid con le sue figure vistose e i suoi titoli a effetto perché non prestiate fede al «National Enquirer»? Ma aspettate un minuto. Questo non è un modo per sfuggire alla domanda? Che processo è se il verdetto di colpevolezza viene usato come parte del capo d'accusa? Quello che ci vuole è un modo per stabilire in modo oggettivo cosa si intende per «vistoso» o «a effetto». La cosa si può dimostrare difficile.

E che dire del rovescio della medaglia? È il formato dignitoso e tradizionale di «Le Scienze», per esempio la sua mancanza di fotografie di celebrità, che vi convince a prestargli fede? Se è così, è un modo decisamente curioso per stabilire cosa sia la verità. Si direbbe che il vostro concetto di verità sia strettamente collegato al vostro modo di valutare lo «stile» di un canale di comunicazione, certamente una nozione piuttosto ineffabile.

Detto questo, devo ammettere che anch'io faccio continuamente affidamento su veloci giudizi di stile nel mio tentativo di scindere il vero dal falso, il credibile dal non credibile. Non saprei dire su quali criteri faccio assegnamento senza averci prima pensato a lungo e aver scritto molte pagine. Ma anche allora, se stessi per pubblicare la guida definitiva *(Come distinguere il vero dal falso mediante il suo stile di pubblicazione)*, essa andrebbe pubblicata per avere qualche effetto, e il suo titolo, per non menzionare lo stile in cui viene pubblicata, attrarrebbe probabilmente alcuni lettori, ma ne respingerebbe molti di più.

Bene, se la verità è così elusiva, non c'è da meravigliarsi che la gente sia assediata da voci contrastanti sulla stampa. Quando ero più giovane, credevo che una volta che qualcosa era stata scoperta, verificata e pubblicata entrasse a far parte della Conoscenza: definitiva e irrevocabile.

Con mia grande sorpresa, però, scoprii che la verità deve continuamente combattere per la propria esistenza. Che un'idea sia stata scoperta e stampata su un «giornale rispettabile» non ci assicura che essa diventi ben nota e accettata. In effetti, di solito dovrà essere riscritta e ristampata molte volte, spesso da persone diverse, prima di avere qualche possibilità di prendere piede. Tutto ciò è sconcertante per un idealista come me, più disposto a credere in una verità monolitica e assoluta che in una verità pluralistica e relativa. Che un'idea debba lottare per la propria esistenza è una scoperta triste. L'idea che la verità *non* verrà fuori, a meno che le sia data una buona spinta, è sconvolgente.

C'è una domanda che si pone in tutte le società: è meglio che tutte le voci differenti lottino tra loro o è meglio avere solo poche pubblicazioni «ufficiali» che dicano ciò che va bene e ciò che non va bene? La nostra società ha optato per una pluralità di voci, per un «mercato delle idee», per un concorso aperto a tutte le teorie in conflitto. Ma se c'è un simile caos, chi ci assicura che ci sia legge e ordine, chi starà a guardia della verità? La risposta è: ci sarà il CSICOP!

CSICOP? Cos'è il CSICOP? È una sigla per Committee for the Scientific Investigation of Claims of the Paranormal. È un nome piuttosto esoterico per un'organizzazione il cui scopo non è affatto esoterico: applicare il buon senso alle affermazioni bizzarre, implausibili e inverosimili.

Chi sono le persone che formano il CSICOP e cosa fanno? L'organizzazione venne fondata da Paul Kurtz, professore di filosofia presso la State University of New York di Buffalo, che le diede vita perché riteneva che fosse necessario contrastare la montante marea di credenze irrazionali e fornire al pubblico una trattazione più bilanciata delle affermazioni relative al paranormale presentando il discordante punto di vista scientifico. Tra i primi membri del CSICOP c'erano alcuni dei più importanti filosofi d'America (per esempio Ernest Nagel e W. V. Quine) e altri pittoreschi combattenti dell'occulto, come lo psicologo Ray Hyman, il mago James Randi e uno che i lettori di questa rubrica hanno forse sentito nominare: Martin Gardner. Nei primi incontri fu deciso che la funzione principale del comitato sarebbe stata la pubblicazione di una rivista dedicata alla sottile arte dello smantellamento. Forse «smantellamento» non è il termine che avrebbero scelto, ma si adatta bene. La rivista che cominciarono a pubblicare nell'autunno del 1976 era «The Zetetic», da una parola greca che significa «ricercatore scettico».

Come avviene per molti movimenti appena nati, si sviluppò una disputa filosofica tra due fazioni, una più «relativistica» e propensa a una sospensione del giudizio, l'altra più rigidamente avversa al nonsenso, più desiderosa di partire all'offensiva e di attaccare le affermazioni relative al soprannaturale. Strano a dirsi, la fazione più aperta non fu così aperta da accettare l'avverso punto di vista e la spaccatura si fece più ampia. Infine ci fu uno scisma. La fazione relativistica (un membro) iniziò a pubblicare un suo giornale, «The Zetetic Scholar», in cui coabitavano felicemente scienza e pseudoscienza. La fazione più ampia mantenne il nome CSICOP e cambiò il titolo del suo giornale in «The Skeptical Inquirer».

L'obiettivo di «The Skeptical Inquirer» è semplicemente quello di combattere il nonsenso. Per riuscirci ricorre al senso comune, il che significa che è accessibile a chiunque sappia leggere. Non c'è bisogno di alcuna particolare conoscenza per leggere le sue pagine, in cui le affermazioni prive di senso sono a una a una fatte a pezzi. (A volte le affermazioni sono chiaramente stupide come i titoli all'inizio di questo articolo, a volte sono molto più sottili.) Tutto quello che si richiede per leggere questo anomalo giornale è la curiosità sulla natura della verità: curiosità su come la verità si difende (attraverso il suo agente CSICOP) dagli attacchi senza quartiere di teorizzatori, pensatori, eccentrici, picchiatelli e semplici truffatori incredibilmente immaginosi.

Il giornale è passato dal piccolo numero di iniziali sottoscrittori a circa 7500 lettori, un Davide se paragonato ai Golia precedentemente ricordati, con la loro circolazione calcolabile in milioni di copie. Le sue pagine sono piene di scritti vivaci e umoristici: la lotta di idee nella sua forma più godibile. Il giornale non è affatto una voce monolitica, l'avvocato di un unico dogma: è piuttosto esso stesso un mercato di idee. Anche le persone che padroneggiano con abilità lo strumento del buon senso possono farlo con stili differenti, e a volte possono essere in disaccordo.

C'è qualcosa di paradossale nelle decisioni editoriali di una simile rivista. Dopo tutto, ciò che è in discussione è essenzialmente la natura delle argomentazioni corrette. Cosa va accettato e cosa no? Per dare un'idea caricaturale della situazione, immaginiamo i dilemmi editoriali di giornali con testate quali «Il bollettino della stampa libera», «Larghe vedute» o «Informazioni di politica editoriale». Quali lettere andrebbero pubblicate, quali articoli? Quale linea andrebbe seguita per scegliere il materiale di redazione?

Non sono domande facili: c'è di mezzo un paradosso, un groviglio in cui le idee da valutare sono anche le idee che compiono la valutazione. L'unica risorsa è il buon senso, la base minima di ogni razionalità. Sfortunatamente non abbiamo a disposizione alcun algoritmo indiscutibile per dare una caratterizzazione univoca a quel profondo strato del razionale, né è verosimile che ci si arrivi presto. Per adesso ci si deve affidare, per giungere al cuore della razionalità, a fattori imperscrutabili come semplicità, eleganza, intuitività. Questo paradosso è sempre esistito in tutto il corso della storia intellettuale, ma nella nostra epoca così ricca di informazioni è particolarmente conturbante.

A dispetto di questi problemi epistemologici, connessi alla sua stessa ragion d'essere, «The Skeptical Inquirer» è fiorente e serve da piacevole antidoto ai dotti giornali scientifici, così spesso curiosamente estranei ai problemi della vita quotidiana. Sotto questo aspetto l'«Inquirer» assomiglia ai rotocalchi scandalistici.

L'elenco degli argomenti dei 17 numeri finora apparsi è estremamente vario. Alcuni argomenti ricorrono solo una volta, altri vengono affrontati regolarmente e discussi sotto varie angolazioni e a vari livelli. Alcuni degli argomenti più ricorrenti sono la percezione extrasensoriale, la telecinesi (l'uso del potere mentale per influenzare eventi a distanza), l'astrologia, i bioritmi, il mostro di Loch Ness, gli UFO, il creazionismo, la telepatia, la visione a distanza, i detective chiaroveggenti che risolvono casi con sicurezza, il triangolo delle Bermude, la «pensierografia» (l'uso del potere mentale per creare immagini su pellicola), la presunta origine extraterrestre della vita sulla Terra, il mago «Don Juan» di Carlos Castaneda, il potere della piramide, la chirurgia psichica e le guarigioni per suggestione, la scientologia, le predizioni di famosi «medium», gli spettri, gli spiriti e le case stregate, la levitazione, la chiromanzia e la lettura della mente, le teorie antropologiche non ortodosse, la sensibilità delle piante, le macchine del moto perpetuo, l'acqua magica e altri tipi di rabdomanzia, le bizzarre mutilazioni del bestiame. E questi non sono gli unici argomenti, sono solo quelli che appaiono regolarmente.

Numerose sono le persone che scrivono di frequente sulla rivista, come Randi, che è veramente prolifico. Tra gli altri ci sono l'esperto in areonautica Philip J. Klass, lo specialista in UFO James E. Oberg, lo scrittore Isaac Asimov, il fondatore del CSICOP (e attuale presidente) Kurtz, lo psicologo James E. Alcock, il pedagogista Elmer Kral, l'antropologo Laurie Godfrey, il giornalista esperto in argomenti scientifici Robert Sheaffer, il sociologo William Sims Bainbridge e molti altri. Anche il direttore della rivista, Kendrick Frazier, un giornalista esperto in argomenti scientifici, pubblica regolarmente eloquenti e mordaci commenti.

Il modo migliore per dare un'idea della rivista è quello di citare dei passi tratti da articoli apparsi sulla rivista stessa. Uno degli articoli che più mi è piaciuto apparve nel secondo numero (primavera/estate 1977). È dello psicologo Ray Hyman (che tra l'altro, come molti altri collaboratori di «The Skeptical Inquirer», è anche un mago di talento) ed è intitolato *Lettura fredda: come convincere gli estranei che sapete tutto su di loro.*

L'articolo inizia prendendo in esame un corso tenuto da Hyman sui vari modi in cui la gente viene manipolata. Racconta Hyman: «Invitai diversi manipolatori di persone perché dessero una dimostrazione delle loro tecniche, imbonitori, venditori di enciclopedie, ipnotizzatori, pubblicitari, predicatori, truffatori e tutta una serie di individui che si occupavano di problemi personali. Le tecniche discusse, soprattutto quelle intese ad aiutare la gente a risolvere i propri problemi personali, sembrano comportare la tendenza del cliente a trovare in ogni situazione più significati di quanti in effetti ce ne siano. Gli studenti accettarono prontamente questa spiegazione quando venne loro proposta, ma non ebbi la sensazione che si rendessero pienamente conto di quanto sia diffusa e radicata questa tendenza umana a trovare senso nel nonsenso.»

Hyman descrive poi il desiderio della gente di credere a quello che gli altri dicono di loro. La sua «regola aurea» è: «Per essere ben accetti dal proprio prossimo, ditegli quello che vuole sentire. Vuole sentir parlare di lui: quindi parlategli di lui. Ma non ditegli quello che sapete di vero su di lui. Oh, no! Non ditegli mai la verità; ditegli invece quello che egli vorrebbe che fosse vero!» Come esempio, Hyman cita il seguente passo (che, per coincidenza, è stato scritto proprio su di *te*, caro lettore!):

«Alcune delle tue aspirazioni tendono a essere davvero irrealistiche. A volte sei estroverso, affabile, socievole, mentre altre volte sei introverso, diffidente e riservato. Hai scoperto che è imprudente essere troppo sinceri nel rivelarsi agli altri. Ti vanti di essere indipendente nelle tue idee e di non accettare le opinioni degli altri senza una prova che ti soddisfi. Preferisci una certa dose di cambiamento e varietà e ti senti insoddisfatto se obbligato a restrizioni e limitazioni. A volte dubiti seriamente di aver preso la giusta decisione o di aver fatto la cosa giusta. Disciplinato e controllato all'esterno, tendi a essere preoccupato e insicuro nel tuo interno.

«Il tuo equilibrio sessuale è stato in qualche modo problematico. Pur avendo alcune debolezze nel carattere, sei generalmente capace di porvi rimedio. Hai molte capacità inutilizzate che non hai volto a tuo vantaggio. Hai una tendenza a criticarti. Hai molto bisogno che gli altri ti apprezzino e ti stimino.»

Azzeccato, eh? Hyman commenta: «Le affermazioni contenute in questo discorsino furono utilizzate per la prima volta nel 1948 da Bertram Forer, che le aveva per lo più tratte da un opuscolo di astrologia comprato in edicola, durante una dimostrazione in classe di autoconferma personale. Gli studenti di Forer, i quali credevano che quel profilo fosse preparato apposta per loro come test della personalità, diedero al profilo un punteggio medio di 4,26 su una scala da 0 (scarso) a 5 (perfetto). Ben 16 su 39 studenti (il 41 per cento) lo qualificarono una perfetta descrizione della loro personalità. Solo cinque diedero un punteggio inferiore a 4 (il punteggio più basso fu 2, cioè "medio"). Quasi 30 anni dopo, degli studenti diedero dello stesso profilo una valutazione quasi identica, considerandolo un'ottima descrizione di se stessi.»

Hyman dà una ricetta in 13 punti per diventare esperti in lettura fredda. Ecco alcuni dei suoi suggerimenti: «Usate la tecnica della "pesca" (far sì che il soggetto parli di sé, poi ripetete quello che ha detto usando altre parole); dare sempre l'impressione di sapere più di quanto dite; non temete di adulare il soggetto ogni volta che ne avete l'occasione». Questa ricetta deliziosamente cinica è presentata in modo molto dettagliato, presumibilmente non per trasformare i lettori dell'articolo in ciarlatani e imbroglioni, ma per mostrar loro come avvengono tali manipolazioni.

Chiede Hyman: «Perché la cosa funziona così bene? Non serve dire che la gente è credulona e suggestionabile; né si può aggirare la domanda sostenendo che alcuni individui non hanno sufficiente capacità di discriminazione o non hanno abbastanza intelligenza per capire l'imbroglio. Al contrario, si può affermare che ci vuole un certo grado d'intelligenza nel cliente perché la lettura funzioni... Dobbiamo far valere la nostra conoscenza e le nostre aspettative perché possiamo comprendere qualcosa del nostro mondo. Nella maggior parte delle situazioni comuni, quest'uso del contesto e della memoria ci rende capaci di interpretare correttamente le informazioni e ci dà i necessari criteri per farlo. Ma questo potente meccanismo può sviarsi in situazioni in cui non c'è alcun effettivo messaggio da veicolare. Invece di afferrare solo un rumore indistinto, noi ci sforziamo di trovare un significato nella situazione. Quindi lo stesso sistema che ci rende capaci di trovare significati in modo creativo e di fare nuove scoperte, ci rende anche estremamente vulnerabili allo sfruttamento operato da ogni sorta di manipolatori. Nel caso della lettura fredda, il manipolatore può essere cosciente di mettere in atto un inganno; ma spesso è egli stesso vittima di un meccanismo di autoconferma.»

(Hyman sa bene quello che dice. Molti anni fa si convinse per un certo periodo di avere egli stesso una vera capacità di leggere il palmo delle mani, finché un giorno provò a dire alla gente l'esatto opposto di quello che le loro mani gli dicevano e scoprì che le sue affermazioni erano più che mai bevute. Cominciò allora a sospettare che la plasticità della mente umana, in particolare la sua, stesse facendo qualcosa di strano.)

All'inizio di ogni numero di «The Skeptical Inquirer» c'è una rubrica intitolata «Novità e commenti» che riguarda le più recenti relazioni su affermazioni sensazionali, gli spettacoli televisivi trasmessi di recente pro o contro il paranormale, i processi di un tipo o l'altro e così via. Una delle cose più divertenti fu il resoconto, pubblicato nel numero dell'autunno 1980, dei «Premi Uri», attribuiti da Randi (il 1° aprile, naturalmente) a persone meritevoli che più di altre avevano contribuito a stimolare creduloneria e credenze irrazionali. Ogni premio consiste di «un cucchiaino d'acciaio inossidabile elegantemente piegato con una base leggera e trasparente». Randi spiegava che i premi erano comunicati per telepatia ai vincitori, i quali erano «liberi, se lo desideravano, di annunciare in anticipo la loro vittoria, mediante la precognizione». I premi erano divisi in quattro categorie: Accademico («allo scienziato che dice le cose più stupide sulla parapsicologia»), Fondazione («all'organizzazione che ricompensa con maggior quantità di denaro le cose più stupide dette a proposito della parapsicologia»), Prestazioni («al medium che, con minor talento, ha il maggior numero di clienti»), e Media («all'organo d'informazione che sostiene le affermazioni più spudorate degli assertori del paranormale»).

La natura delle coincidenze è un tema ricorrente nelle discussioni sul paranormale. Ho un vivo ricordo di un passo del delizioso libro di Warren Weaver *Lady Luck: The Theory of Probability*, in cui egli sottolinea come in molte situazioni l'esito più probabile sia un evento molto improbabile (come nella distribuzione delle carte nel bridge, quando qualsiasi mano abbiate è certamente straordinariamente rara). Un'affermazione analoga è contenuta nel seguente brano tratto da un recente libro di David Marks e Richard Kammann intitolato *The Psychology of the Psychic* (molti passi del quale sono stati ristampati in un numero di «The Skeptical Inquirer»):

«Primo, noi notiamo e ricordiamo le coincidenze, soprattutto le *strane coincidenze*, ogni volta che si verificano. Secondo, *non* notiamo le non-coincidenze. Terzo, il fatto che non notiamo i non-eventi crea l'*illusione di breve durata* che fa sì che la strana coincidenza sembri improbabile. Quarto, abbiamo difficoltà nello stimare le combinazioni di eventi. Quinto, trascuriamo il *principio delle strane coincidenze equivalenti*, secondo il quale una coincidenza vale l'altra per quanto riguarda la teoria medianica.»

Un ottimo esempio di mancata individuazione dei non-eventi è dato dalle previsione sbagliate di famosi indovini (come Jeane Dixon). Per lo più la gente non va mai a vedere se gli eventi hanno confermato le predizioni. «The Skeptical Inquirer», invece, ha l'abitudine di tornare indietro per fare delle verifiche. Alla fine di ogni anno stampa un certo numero di predizioni di vari indovini e valuta il loro successo. Nel numero dell'autunno 1980, i redattori presero le predizioni di 100 indovini di fama, le tabularono, elencarono le prime 12 in ordine di frequenza e lasciarono ai lettori il compito di valutare l'esattezza delle visioni medianiche del futuro. La predizione n. 1 per il 1979 (fatta da 86 indovini) era «Si allungherà il corso della vita in quanto verrà messo sotto controllo il processo di invecchiamento». La n. 2 (85 indovini) era «Ci sarà un importante progresso nella lotta contro il cancro, che verrà quasi totalmente debellato». La n. 3 (ancora 85) era «Ci sarà una sorprendente rinascita spirituale e un ritorno agli antichi valori». Così sono andate le cose. La n. 6 (81 indovini) era «Ci sarà un contatto con alieni venuti dallo spazio che ci daranno un'incredibile quantità di conoscenze».

C'è qualcosa di patetico, anzi di disperato, in queste predizioni. È fin troppo chiara l'analogia tra le pubblicazioni che riportano queste predizioni e i vacui spettacoli televisivi quali «Fantasy Island» e «StarTrek». Il denominatore comune è la fuga dalla realtà. In un articolo del numero dell'autunno 1979 di «The Skeptical Inquirer», William Sims Bainbridge puntualizza bene la questione. Forse abbiamo tutti il desiderio di diluire la realtà con la fantasia, di far sì che la realtà sembri più semplice e più allineata con ciò che noi vorremmo essa fosse. Eppure, allo stesso tempo, forse tutti noi abbiamo la capacità potenziale e perfino il desiderio di separare il senso dal non senso; basterebbe che fossimo portati alla distinzione in modo sufficientemente coercitivo.

Ma come si può fare? Nella rubrica «Novità e commenti» del numero della primavera 1980 di «The Skeptical Inquirer» c'era un passo in cui si parlava di una briosa conferenza di antipseudoscienza tenuta da «Capitan raggio di luce», in realtà Douglas F. Stalker, professore associato di filosofia all'Università del Delaware. L'articolo cita il seguente passo del «comico smantellamento» effettuato da Stalker (contro l'astrologia, i bioritmi, la numerologia, gli UFO, il potere della piramide, le pretese medianiche e simili):

«Per anni ho tenuto serie lezioni contro di loro, caricando in modo diretto le loro stupide teorie. Questi attacchi diretti non hanno fatto breccia in molte menti, quindi ho deciso di seguire un approccio indiretto. Se non riesci a batterli, unisciti a loro; e così ho fatto, per modo di dire. Ho costruito per conto mio delle pseudoscienze chiaramente assurde e ho dimostrato che erano proprio come l'astrologia e le altre. Lavorando dall'interno, un maggior

numero di studenti è arrivato ad accorgersi di quanto pseudo siano queste pseudoscienze... E questo è il pubblico che io cerco di raggiungere: le nuove leve di cittadini. La mia dimostrazione li raggiunge nel modo giusto; lascia un'impressione duratura; procura amici e cambia le menti.» (Sono lieto di riferire che Stalker accetta nuove iscrizioni. Ci si può mettere in contatto con lui presso il Department of Philosophy, University of Delaware, Newark, Del. 1971 l.)

Uno dei rilievi fatti da Stalker è che per quanto eloquente sia una lezione, essa non ha il potere di convinzione dell'esperienza. La stessa osservazione è riportata nel classico studio di Barry Singer e Victor A. Benassi del dipartimento di psicologia della California State University di Long Beach. Questi due ricercatori si erano prefissi di stabilire l'effetto su studenti del primo anno di psicologia provocato da fenomeni apparentemente paranormali creati in classe da un mago vestito in modo esotico. I loro risultati furono riportati nel numero dell'inverno 1980/81 di «The Skeptical Inquirer» in un articolo intitolato *Fooling Some of the People All of the Time*.

In due delle classi, il mago (Craig Reynolds) fu presentato come un laureando «interessato alla psicologia del paranormale ovvero alle capacità medianiche, che aveva lavorato al raffinamento di una presentazione delle sue capacità medianiche». L'istruttore affermò anche esplicitamente: «Io, personalmente, non sono convinto delle capacità medianiche di Craig né di alcun altro». In altre due classi Craig fu presentato come un laureando «interessato alla psicologia del magico e dei trucchi da palcoscenico, che aveva elaborato una presentazione del suo spettacolo di magia». Fu sottolineato che tutti i numeri di Craig erano «semplici trucchetti da dilettante praticati da secoli».

Dopo lo spettacolo, agli studenti fu chiesto di fare una relazione sulle loro reazioni. Singer e Benassi ricevettero dure scosse dalla lettura delle relazioni: «Primo... sia nelle classi "magiche" che in quelle "medianiche", circa due terzi degli studenti erano chiaramente convinti che Craig fosse un individuo dotato di poteri occulti. Solo pochi studenti sembravano credere all'affermazione dell'istruttore che Craig era un mago, nelle classi in cui era stato presentato come tale. Secondo, la convinzione medianica non solo era prevalente; era anche forte e carica di emozione. Un certo numero di studenti coprì i fogli di esorcismi e scongiuri contro il diavolo. Nella situazione medianica, il 18 per cento degli studenti espresse esplicitamente paura e disturbi emotivi. La maggior parte degli studenti espresse timore e stupore.

«Noi eravamo presenti a due delle esibizioni di Craig e fummo testimoni di comportamenti abnormi. Mentre Craig era impegnato nella cantilena "del piegamento" (parte di un numero in cui piegava una sbarra di acciaio), la classe era terribilmente eccitata. Gli studenti sedeva rigidamente sulle sedie, occhi sbarrati e bocca aperta, cantilenando insieme. Quando la sbarra si piegò, l'aula si riempì di sospiri e mormorii. Finita l'esibizione e congedata la classe, gli studenti rimasero seduti a guardare nel vuoto o a scuotere la testa, oppure si precipitarono eccitati verso Craig, chiedendogli come avrebbero potuto fare per sviluppare simili poteri. Ci accorgemmo di stare osservando un effetto comportamentale di straordinario vigore. Se Craig, alla fine della sua esibizione, avesse chiesto agli studenti di togliersi i vestiti, dargli i loro soldi e dar vita a un nuovo culto, crediamo che alcuni avrebbero aderito con entusiasmo. Stava ovviamente avvenendo qualcosa che non riuscivamo a capire.»

Dopo questa teatrale presentazione, alle classi fu detto che quello che avevano visto erano solo dei trucchi. Va anche detto che altre due classi avevano assistito alla medesima rappresentazione, ma con questo avvertimento: «Nella sua esibizione Craig pretenderà di leggere nelle menti e di dimostrare capacità medianiche, ma Craig non ha per davvero simili capacità e quello che vedrete sono solo trucchi». Nonostante questa premessa, più di metà degli studenti di queste classi si convinsero che egli fosse dotato di poteri occulti. Osservano filosoficamente Singer e Benassi: «Tutto ciò è significativo o in rapporto allo status degli insegnanti universitari davanti ai loro studenti o per quanto riguarda gli strani percorsi che la gente segue per giungere a credere nell'occulto.»

Quello che segue è veramente sorprendente: «La domanda successiva chiedeva se i prestigiatori saprebbero fare tutto quello che aveva fatto Craig, e in teoria tutti gli studenti erano concordi nel dare una risposta affermativa. Fu allora chiesto se avrebbero voluto rivedere il loro giudizio sulle capacità medianiche di Craig alla luce di questa informazione negativa che loro stessi aveva fornito. Solo pochi lo fecero, riducendo la percentuale di studenti convinti dei poteri occulti di Craig al 55 per cento.

«Fu poi chiesto agli studenti di valutare quanti, tra coloro che erano in grado di riprodurre i numeri effettuati da Craig, usavano in realtà trucchi da prestigiatore. La risposta generale fu che almeno tre "medium" su quattro erano in realtà imbroglioni. Dopo che ebbero fornito questa informazione negativa, fu nuovamente chiesto loro se volessero rivedere il loro giudizio sui poteri di Craig: di nuovo, solo pochi lo fecero e la percentuale di coloro che credevano ai poteri occulti di Craig si ridusse al 52 per cento.»

Singer e Benassi commentano: «Cosa ci dice tutto ciò? I risultati del nostro test scritto suggeriscono come la gente possa continuare a credere ingenuamente nei poteri occulti di qualcuno *anche quando ne sa di più*. È un errore logico continuare a ritenere che un individuo (Craig in questo caso) sia un medium, pur ammettendo nello stesso tempo che dei prestigiatori saprebbero fare le stesse cose che fanno dei veri medium e giudicando che la maggior parte dei cosiddetti medium sono degli imbroglioni. È proprio vero che noi uomini siamo così stupidi? Ebbene sì.»

Pochi anni fa, Scot Morris (ora senior editor di «Omni» e responsabile della rubrica «Giochi») compì un esperimento analogo in una classe del primo anno di psicologia della Southern Illinois University e pubblicò i risultati sul numero della primavera 1980 di «The Skeptical Inquirer». Innanzitutto Morris valutò le convinzioni degli studenti a proposito della percezione extrasensoriale facendo loro compilare un questionario. Poi un suo collega effettuò una «dimostrazione ESP», che Morris giudica «spaventosamente impressionante».

Dopo questa esibizione, Morris cercò di «deprogrammare» i suoi studenti. Aveva due armi a sua disposizione. Una è ciò che egli chiama «smascheramento», un'operazione di tre minuti che consisteva nel rivelare come funzionavano due dei tre trucchi, confessando che anche l'altro dei tre sconcertanti numeri era un trucco; «ma», diceva Morris, «non vi dirò com'era perché voglio che vi rendiate conto di come non sia necessario considerare soprannaturale ciò che non si riesce a spiegare». L'altra arma era una lezione anti-ESP di 50 minuti in cui venivano rivelati i segreti dei lettori della mente professionisti, erano analizzate le stime comuni delle probabilità delle «strane coincidenze», si dimostrava come gli studi «scientifici» sulla percezione extrasensoriale fossero discutibili per varie ragioni statistiche e logiche e venivano addotte altre semplici ragioni per mettere in serio dubbio l'esistenza della ESP.

Dopo l'esibizione, solo metà degli studenti furono informati del trucco, mentre tutti ascoltarono la lezione anti-ESP. Fu poi fatta un'inchiesta sulla forza delle loro convinzioni relative a vari tipi di fenomeni paranormali e risultò che gli studenti informati del trucco credevano nella ESP molto meno di quelli che avevano soltanto ascoltato la lezione. Il livello medio di accettazione della ESP tra gli studenti informati del trucco scese da circa 6 (moderata accettazione) a circa 2 (forte scetticismo), mentre il livello medio tra gli studenti non informati scese da 6 a circa 4 (lieve scetticismo). Dice Morris facendo il punto sui risultati: «l'esperienza di smascheramento fu chiaramente cruciale; per produrre scetticismo nei confronti della ESP fu più efficace la rivelazione in tre minuti che erano stati presi in giro della lunga denuncia della ESP durata un'ora».

Tra gli interessi che avevano spinto Morris a condurre l'esperimento c'era il desiderio di controllare «se l'esercizio avrebbe spinto gli studenti a essere più scettici solo nei confronti delle affermazioni sulla ESP oppure, come speravamo, avrebbe indotto un più generale atteggiamento di scetticismo. Per esempio, la loro esperienza li avrebbe resi più scettici sull'astrologia, le sedute spiritiche e i fantasmi?» Morris trovò un moderato trasferimento di scetticismo e concluse fiduciosamente che «insegnare a qualcuno a essere scettico a proposito di una credenza, lo rende più scettico a proposito di cre-

denze simili e forse un po' più scettico anche a proposito di credenze dissimili».

Io credo che la questione del trasferimento di scetticismo sia fondamentale. Serve poco imparare una lezione se quella lezione rimane su dei casi particolari e non è applicabile al di là del caso in cui è stata appresa la prima volta. Qual era la lezione, per esempio, del dramma di Jonestown? Semplicemente che non si dovrebbe mai seguire Jim Jones in Guyana? O, più generalmente, che sarebbe bene evitare di seguire un guru per il mondo? O che non si dovrebbe mai seguire nessuno in alcun luogo? O che tutti i culti sono demoniaci? O che credere in qualsiasi tipo di sapere, umano o divino, è folle e pericoloso? Perché no? A quale livello di generalizzazione si apprende una lezione?

Il Capitan raggio di luce di Stalker era convinto, smantellando le sue stesse pseudoscienze «in miniatura» davanti a un pubblico, di poter trasferire alla gente una più generale capacità critica, una capacità di pensare in modo più chiaro alle pretese relative al paranormale. È una convinzione fondata? Ci sono molti assertori di certi tipi di fenomeni paranormali che ne ridicolizzano altri tipi. C'è gente che si fa beffe dei titoli del «National Enquirer» e nello stesso tempo crede, ad esempio, che si possa imparare a levitare con la meditazione trascendentale, o che si verifichino le predizioni astrologiche, o che gli UFO siano visitatori d'altri mondi, o che esista la percezione extrasensoriale. Molti hanno detto: «Sfortunatamente, la maggior parte dei medium sono degli imbroglioni, il che rende più difficile riconoscere quelli *autentici*». C'è perfino chi crede in gente come Uri Geller e dice: «Ammetto che alle volte imbrogli, forse anche il 90 per cento delle volte, ma continuo a credere che abbia davvero delle capacità paranormali.»

Se uno cerca un segnale in mezzo al rumore, e più cerca e più rumore trova, quando è ragionevole che rinunci e concluda che non c'è alcun segnale? D'altra parte, a volte potrebbe davvero esserci un segnale. Il problema è che non si vuole saltare troppo rapidamente a una generalizzazione negativa, soprattutto se le proprie opinioni sono influenzate dal contesto. In fondo, non tutto quello che pubblica il «National Enquirer» è falso. Il difficile è valutare quando le prove sono sufficienti e quando non lo sono. Se poi si tratta di un bene o di un male, è un fatto soggettivo, e un problema che finora pochi giornali hanno trattato.

«The Skeptical Inquirer» si occupa di questioni che vanno dal ridicolo al sublime, dal banale al profondo. C'è chi potrebbe dire che sia una gran perdita di tempo preoccuparsi di sciocchezze come la ESP o altri cosiddetti effetti paranormali. Altri, e io sono uno di loro, ritengono che chiunque non sia capace o non voglia impegnarsi a riflettere su ciò che distingue il sistema di pensiero scientifico dai suoi molti sistemi rivali non è affatto un sostenitore della verità, e che inoltre il diffondersi del nonsenso è una tendenza pericolosa che va tenuta sotto controllo.

C'è sempre il problema, comunque, di vedere se «The Skeptical Inquirer» riuscirà mai a essere qualcosa di più di una goccia nel mare. Certo i suoi redattori non si aspettano di vederlo un giorno in vendita accanto al «National Enquirer» alle casse dei supermarket. E si può immaginare uno scenario rovesciato in cui un giornale d'opposizione come «The Skeptical Inquirer» vende milioni di copie nei supermarket, insieme ai molti rivali, mentre una coraggiosa voce isolata dell'occulto esce quattro volte l'anno in formato tabloid ed è letto da 7500 persone? Sembra un'assurda storia, buona per le pagine del «National Enquirer»! Questo ridicolo scenario serve a chiarire meglio la situazione in cui si trova a operare il coraggioso gruppo del CSICOP.

A che serve pubblicare un giornale che viene letto solo da una piccola schiera di già convinti oppositori dell'occulto? La risposta si trova, tra l'altro, nella rubrica delle lettere che chiude ogni numero. Molte persone scrivono per dire quanto importante sia stata la rivista per loro, i loro amici e i loro studenti. Gli insegnanti delle superiori sono tra quelli che più di frequente scrivono lettere di ringraziamento alla rivista, ma abbiamo visto lettere entusiastiche anche di membri del clero, di conduttori di rubriche radiofoniche e di rappresentanti di altre professioni.

Spero di aver sollevato abbastanza interesse tra i lettori perché alcuni di loro vogliano abbonarsi a «The Skeptical Inquirer». La quota è di 16 dollari all'anno e l'indirizzo è Box 229, Central Park Station, Buffalo, N.Y. 14215. Per i lettori di ampie vedute do anche l'indirizzo e il costo dell'abbonamento di «The Zetetic Scholar» (Department of Sociology, Eastern Michigan University, Ypsilanti, Mich. 48197, 12 dollari all'anno) e del «National Enquirer» (Lantana, Fla. 33464, 13,95 dollari all'anno).

Certo nessuno riuscirà mai a colmare il vasto oceano dell'irrazionalità che ci circonda, ma l'ambizione di «The Skeptical Inquirer» non è mai stata così grande; la sua ambizione, piuttosto, è quella di essere una solida boa a cui aggrapparsi in quel mare tempestoso, e anche quella di stimolare una sana vena di scetticismo nel maggior numero di persone possibile. Come disse Frazier in uno dei suoi eloquenti editoriali: «Lo scetticismo non è, come molti ritengono, un punto di vista. È invece una componente essenziale della ricerca intellettuale, un metodo per determinare i fatti qualsiasi possano essere e ovunque possano condurre. Fa parte di ciò che chiamiamo senso comune ed è uno dei modi in cui la scienza lavora».

È un vero peccato dover continuamente difendere la verità contro tanti attacchi condotti da gente che si rifiuta di pensare, ma d'altronde la sciatteria di pensiero sembra inevitabile, fa parte della natura umana. A proposito, mi pare di aver letto di recente da qualche parte che l'individuo tipo utilizza solo il 10 per cento del suo cervello. Altro che sciatteria: è davvero sorprendente! Perfino gli scienziati sono sconcertati!